Planning 8/3/80

1. Sequence Rosa ... Huly Chris Mo
Hin I Rik ?
Min II
Min II
ombre

This is EXHIBIT FIERS-14
to
the Affidavit of Walter C. Fiers
sworn before me
this 19 th day of November, 2001

Commissioner for Oath or Notary Public

2. Expenses
a) Samples in Leuven : cl 3 S100
+ antiserum F-IF
+ antiserum L-IF
+ trypsine
+ controle [illegible]

b) Samples naar Leuven :
Kalinda i Bact. + F-IF open → S100
ii Bact open → L-IF → S100
iii Bact S100 + F-IF
iv 10 min 37 °C
v Buffer + FIF

cl 3 S100
i ... antiserum | als kalibrering serie met antiserum + controles [illegible]
standaard. F-IF + antiserum
ii ... trypsine | serie met trypsine — als kalibrering — trypsine inhibitor
iii. bacterial extract + F-IF
iii ... inhibitor ... buffer pH2
iv Warmte-inhibitor (verzijden ...)

Opened / Déposé le Dec 20 20 02
Commissioner of Patents / Commissaire des brevets
In presence of examiner / en présence de l'examinateur

v Extract maken met een kleine hoeveelheid NP40 ..
vi Koncentratie
a) CPG. (controlled pored glas poeder) (→ opgeloste membranen [illegible])
b) Am_2SO_4 ..... + dialyse.
c) Drukdialyse → 1/5 – (Amicon ?).
d 7 id + antiserum

c) Expressie formulen
(i) No Bal :
dag 3 – 4 niet induceerd
induceerd [illegible] 1x
2x id + antiserum
3x

controle anti-sense
(ii) Bal treatment : serie
(afhankelijk v clone analyse).

d) $2'5$ A synthetase assay : geneest volume 5 ml.
d. 3
5K d 3 – 4 | [illegible]
5K . (antisense) |
5K.K [illegible] |

e) [illegible] immunoprecipitatie :
zelfde reeks als onder d) een lage specifieke meting

[illegible] J. [illegible]

25-03-80

Linear DNA

Acc I sites in INTERFERON

GTAGAC (c2)
0
GTATAC (c2)
0
GTCGAC (c2)
0
GTCTAC (c2)
0

Resulting fragment sizes :
850
Tabled according to length :
850

Atu BI sites in INTERFERON

CCAGG (c2)
462
CCTGG (c2)
388 430 552

Resulting fragment sizes :
389 42 32 90 297
Tabled according to length :
389 297 90 42 32

Acy I sites in INTERFERON

GGCGCC (c2)
0
GGCGTC (c2)
0
GACGCC (c2)
287
GACGTC (c2)
0

Resulting fragment sizes :
288 562
Tabled according to length :
562 288

Asu I sites in INTERFERON

GGGCC (c1)
0

```
GGACC           (c1)
  555
GGTCC           (c1)
    0
GGCCC           (c1)
    0

Resulting fragment sizes :
  555  295
Tabled according to length :
  555  295


                Ava I sites in INTERFERON

CCCGGG          (c1)
    0
CTCGGG          (c1)
    0
CCCGAG          (c1)
    0
CTCGAG          (c1)
    0

Resulting fragment sizes :
  850
Tabled according to length :
  850


                Ava II sites in INTERFERON

GGACC           (c1)
  555
GGTCC           (c1)
    0

Resulting fragment sizes :
  555  295
Tabled according to length :
  555  295


                Ava III sites in INTERFERON

ATGCAT          (cX0)
    0

Resulting fragment sizes :
  850
Tabled according to length :
  850


                Avr II sites in INTERFERON

CCTAGG          (cX0)
    0
```

```
Resulting fragment sizes :
  850
Tabled according to length :
  850
```

Alu I sites in INTERFERON

```
AGCT            (c2)
  118   131   183   264

Resulting fragment sizes :
  119    13    52    81   585
Tabled according to length :
  585   119    81    52    13
```

Bam HI sites in INTERFERON

```
GGATCC          (c1)
    0

Resulting fragment sizes :
  850
Tabled according to length :
  850
```

Bbv I sites in INTERFERON

```
GCTGC           (cX0)
  265
GCAGC           (cX0)
  162   262   268

Resulting fragment sizes :
  161   100     3     3   583
Tabled according to length :
  583   161   100     3     3
```

Bcl I sites in INTERFERON

```
TGATCA          (c1)
    0

Resulting fragment sizes :
  850
Tabled according to length :
  850
```

Bgl II sites in INTERFERON

```
AGATCT          (c1)
  629
```

```
 Resulting fragment sizes :
   629  221
 Tabled according to length :
   629  221


                  Bst EII sites in INTERFERON

 GGTGACC           (c1)
     0
 GGTAACC           (c1)
     0
 GGTCACC           (c1)
     0
 GGTTACC           (c1)
   611

 Resulting fragment sizes :
   611  239
 Tabled according to length :
   611  239


                  Bal I sites in INTERFERON

 TGGCCA            (c3)
     0

 Resulting fragment sizes :
   850
 Tabled according to length :
   850


                  Cla I sites in INTERFERON

 ATCGAT            (c2)
     0

 Resulting fragment sizes :
   850
 Tabled according to length :
   850


                  Sau II sites in INTERFERON

 CCGGG            (c2)
     0
 CCCGG            (c2)
     0

 Resulting fragment sizes :
   850
```

```
 Resulting fragment sizes :
   850
 Tabled according to length :
   850


                  Pvu II sites in INTERFERON

 CAGCTG           (c3)
   263

 Resulting fragment sizes :
   265  585
 Tabled according to length :
   585  265


                  Pst I sites in INTERFERON

 CTGCAG           (c5)
   266

 Resulting fragment sizes :
   270  580
 Tabled according to length :
   580  270


                  Rsa I sites in INTERFERON

 GTAC           (c2)
   533  717

 Resulting fragment sizes :
   539  179  132
 Tabled according to length :
   539  179  132


                  Sma I sites in INTERFERON

 CCCGGG           (c4)
     0

 Resulting fragment sizes :
   850
 Tabled according to length :
   850


                  Sfa NI sites in INTERFERON

 GATGC          (cX0)
   310  639
```

```
GCATC          (cX0)
    0

Resulting fragment sizes :
  309  379  162
Tabled according to length :
  379  309  162
```

Sac I sites in INTERFERON

```
GAGCTC         (c5)
    0

Resulting fragment sizes :
  850
Tabled according to length :
  850
```

Sac II sites in INTERFERON

```
CCGCGG         (c4)
    0

Resulting fragment sizes :
  850
Tabled according to length :
  850
```

Sac III sites in INTERFERON

```
ACGT           (cX0)
    0

Resulting fragment sizes :
  850
Tabled according to length :
  850
```

Sal I sites in INTERFERON

```
GTCGAC         (c1)
    0

Resulting fragment sizes :
  850
Tabled according to length :
  850
```

Sph I sites in INTERFERON

```
GCATGC         (c5)
    0
```

```
 Resulting fragment sizes :
   850
 Tabled according to length :
   850


                  Taq I sites in INTERFERON

TCGA           (cl)
     9

 Resulting fragment sizes :
     9  841
 Tabled according to length :
   841    9


                  Xba I sites in INTERFERON

TCTAGA         (cl)
     0

 Resulting fragment sizes :
   850
 Tabled according to length :
   850


                  Xho I sites in INTERFERON

CTCGAG         (cl)
     0

 Resulting fragment sizes :
   850
 Tabled according to length :
   850


                  Xho II sites in INTERFERON

AGATCC         (cl)
     0
AGATCT         (cl)
   629
GGATCC         (cl)
     0
GGATCT         (cl)
     0

 Resulting fragment sizes :
   629  221
 Tabled according to length :
   629  221
```

Xma I sites in INTERFERON

CCCGGG (c1)
0

Resulting fragment sizes :
850
Tabled according to length :
850

Xma III sites in INTERFERON

CGGCCG (c1)
0

Resulting fragment sizes :
850
Tabled according to length :
850

Bgl I sites in INTERFERON

GCCNNNNNGGC (c7)
0

Tth111 I sites in INTERFERON

GACNNNGTC (c4)
0

Ecc e sites in INTERFERON

~~TCANNNNNNNNNGTC~~ ~~(cX0)~~
0

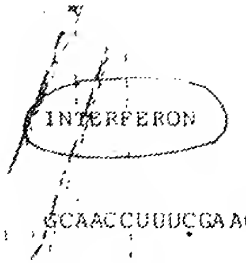
TGANNNNNNNNNTGCT (cX0)
0

AGCANNNNNNNNNTCA (cX0)
0

Ecc K sites in INTERFERON

AACNNNNNNGTGC (cX0)
0

GCACNNNNNNGTT (cX0)
0



MetThrAsnLysCysLeuLeuGlnIleAlaLeuLeu

GCAACCUUUCGAAGCCUUUGCUCUGGCACAACAGGUAGUAGGCGACACUGUUCGUGUUGUCAACAUGACCAACAAGUGUCUCCUCCAAAUUGCUCUCCUG 100

LeuCysPheSerThrThrAlaLeuSerMetSerTyrAsnLeuLeuGlyPheLeuGlnArgSerSerAsnPheGlnCysGlnLysLeuLeuTrpGlnLeuA

UUGUGCUUCUCCACUACAGCUCUUUCCAUGAGCUACAACUUGCUUGGAUUCCUACAAAGAAGCAGCAAUUUUCAGUGUCAGAAGCUCCUGUGGCAAUUGA 200

snGlyArgLeuGluTyrCysLeuLysAspArgMetAsnPheAspIleProGluGluIleLysGlnLeuGlnGlnPheGlnLysGluAspAlaAlaLeuTh

AUGGGAGGCUUGAAUACUGCCUCAAGGACAGGAUGAACUUUGACAUCCCUGAGGAGAUUAAGCAGCUGCAGCAGUUCCAGAAGGAGGACGCCGCAUUGAC 300

rIleTyrGluMetLeuGlnAsnIlePheAlaIlePheArgGlnAspSerSerSerThrGlyTrpAsnGluThrIleValGluAsnLeuLeuAlaAsnVal

CAUCUAUGAGAUGCUCCAGAACAUCUUUGCUAUUUUCAGACAAGAUUCAUCUAGCACUGGCUGGAAUGAGACUAUUGUUGAGAACCUCCUGGCUAAUGUC 400

TyrHisGlnIleAsnHisLeuLysThrValLeuGluGluLysLeuGluLysGluAspPheThrArgGlyLysLeuMetSerSerLeuHisLeuLysArgT

UAUCAUCAGAUAAACCAUCUGAAGACAGUCCUGGAAGAAAAACUGGAGAAAGAAGAUUUCACCAGGGGAAAACUCAUGAGCAGUCUGCACCUGAAAAGAU 500

yrTyrGlyArgIleLeuHisTyrLeuLysAlaLysGluTyrSerHisCysAlaTrpThrIleValArgValLysIleLeuArgAsnPheTyrPheIleAs

AUUAUGGGAGGAUUCUGCAUUACCUGAAGGCCAAGGAGUACAGUCACUGUGCCUGGACCAUAGUCAGAGUGAAAAUCCUAAGGAACUUUUACUUCAUUAA 600

nArgLeuThrGlyTyrLeuArgAsn

CAGACUUACAGGUUACCUCCGAAACUGAAGAUCUCCUAGCCUGUGCCUCUGGGACUGGACAAUUGCUUCAAGCAUUCUUCAACCAGCAGAUGCUGUUUAA 700

GUGACUGAUGGCUAAUGUACUGCAUAUGAAAGGACACUAGAAGAUUUUGAAAUUUUUAUUAAAUUAUGAGUUAUUUUUAUUUAUUUAAAUUUUAUUUL3G 800

AAAAUAAAUUAUUUUUGGUGCAAAAGUCAAAAAAAAAAAAAAAAAAAAA

15

INTERFERON

Opened / Décacheté — Dec 30 02

In presence of examiner / en présence de l'examinateur

	5
	12
	19
	26
	2
	9
	16
	23
	30
	6
	13
	20
	27
	4
	11
	18
	25
	1
	8
	15
	22
	29
	6
	13
	20
	27

1980

DAILY REMINDER

THE STANDARD DIARY DIVISION
OF
WILSON JONES COMPANY

To remember this diary use number stamped on the outer binding edge of cover

This is EXHIBIT FIERS-16
to
the Affidavit of Walter C. Fiers
sworn before me
this 15 th day of November, 2001

Commissioner for Oath or Notary Public

TUESDAY 25 MARCH

follow 1980 85th day – 281 days follow

-- 10^{30} – 11^{30} (8? (Fiori) telephone
call Fiori / [illegible]
EAJ/ EG

MINUTES OF SCIENTIFIC BOARD MEETING
on March 28-29th, 1980

Friday, March 28th :

This is EXHIBIT FIERS-17
to
the Affidavit of Walter C. Fiers
sworn before me
this 13 th day of November, 2001

Commissioner for Oath or Notary Public

Fibroblast interferon

W. Fiers reports on the successful cloning of cDNA genes for fibroblast interferon.

An unsuspected finding was the presence in most clones of an additional sequence of inverted polarity.

The entire nucleotide and amino acid sequence was determined, from the study of 2 clones. It is planned to reconstruct one full gene from those 2 clones via a PST site.

A word of caution is presented on possible artefacts in assay for activity (induction of interferon synthesis by bacterial extracts).

$O_{1/4}$	1.2	0
	1.2	0
	1.0	0*
	1.2	0**
$O_{1/5}$	0.7	0
	0.7	≤0.2*
	1.0	0**
$O_{1/6}$	0.7	0
	1.0	≤0.2*
	0.5	0**
$O_{1/7}$	0.5	0
	1.2	0*
	<0.2	0.5**
$O_{1/8}$	0	1.7*
	<0.2	1.2*
	0	0.7**
	0	1.0**

---

* DBM cellulose paper method

** Nitrocellulose sheets

Therefore, clone $O_{1/8}$ contains a recombinant DNA molecule capable of hybridizing F IF mRNA from total RNA containing F IF mRNA. Non-specific RNA·DNA binding is highly unlikely, because a comparison of Fractions 1A and 4A revealed substantially no non-specific binding of STNV DNA in these same experiments. E.g. as monitored by translation in a rabbit reticulocyte lysate in the presence of ^{35}S-methionine, followed by gel electrophoresis, as described above. Clone $O_{1/8}$ was designated E. coli HB101 (G-pBR322(Pst)/HFIF1 ("G-HB101-pHFIF1"), its recombinant DNA molecule G-pBR322 (Pst)HcFIF1 ("pHFIF1") and its hybrid insert "pHFIF1 fragment". This nomenclature indicates that the clone and recombinant DNA molecule originated in Ghent ("G") and comprises plasmid pBR322 containing, at the PstI site HFIF cDNA ("HFIF"), the particular molecule being the first located.

This is EXHIBIT FIERS-18
to
the Affidavit of Walter C. Fiers
sworn before me
this 19 th day of November, 2001

IDENTIFICATION OF CLONES CONTAINING RECOMBINANT DNA-MOLECULES CROSS-HYBRIDIZING TO pHFIF1

pHFIF1, isolated above, was used to screen the library of clones, prepared previously, for bacterial clones containing recombinant DNA molecules having related hybrid DNA inserts, by colony hybridization (M. Grunstein and D.S. Hogness, "A Method For The Isolation Of Cloned DNA's That Contain A Specific Gene", Proc. Natl. Acad. Sci. USA, 72, pp. 3961-3965 (1975)). This method allows rapid identification of related clones by hybridization of a radioactive probe to the DNA of lysed bacterial colonies fixed in nitrocellulose filters.

The library of clones stored in microtiter plates as described above, was replicated on similar size nitrocellulose sheets (0.45 um pore-diameter, Schleicher and Schuel or Millipore), which had been previously boiled to remove detergent, and the sheets placed on LB-agar plates, containing tetracycline at 10 ug/ml. Bacterial colonies were grown overnight at 37°C. Lysis and fixation of the bacteria on the nitrocellulose sheets took place by washing consecutively in 0.5 N NaOH (twice for 7 min), 1 M Tris-HCl (pH 7.5)(7 min), 0.5 M Tris-HCl (pH 7.5) and 1.5 M NaCl (7 min), 2 x SSC (0.15 M NaCl, 0.015 M sodium citrate (pH 7.2)(for 7 min)). After thorough rinsing with ethanol and air drying, the sheets were baked at 80°C for 2 h in vacuo and stored at room temperature.

A Hinf I restriction fragment specific for the pHFIF1 fragment (infra) served as the probe for colony hybridization, described infra. This fragment (~170 base-pairs) was purified by electrophoresis of the Hinf digestion products of pHFIF1 in a 6% polyacrylamide gel. After staining the DNA bands with ethidiumbromide, the specific fragment was eluted, reelectrophoresed and ^{32}P-labelled by "nick translation" (P.W.J. Rigby et al., "Labeling Deoxyribonucleic Acid To High Specific Activity In Vitro By Nick Translation With DNA Polymerase I", J. Mol. Biol.,

113, pp. 237-251 (1977)) by incubation in 50 µl 50 mM Tris-HCl (pH 7.4), 10 mM $MgCl_2$, 20 mM β-mercaptoethanol, containing 2.5 µl each of dCTP, dTTP and dGTP at 400 µM, 100 pmoles α-^{32}P-ATP (Amersham, 2000 Ci/mmole) and 2.5 units of DNA-polymerase I (Boehringer) at 14°C for 45 min. The unreacted deoxynucleoside triphosphates were removed by gel filtration over Sephadex G-50 in T.E. buffer. The highly ^{32}P-labelled DNA was precipitated with 0.1 vol of 2 M sodium acetate (pH 5.1) and 2.5 vol of ethanol at -20°C.

Hybridization of the above probe to the filter impregnated DNA was carried out essentially as described by D. Hanahan and M. Meselson (personal communication): The filters, prepared above, were preincubated for 2 h at 68°C in 0.1% Ficoll, 0.1% polyvinylpyrrolidone, 0.1% bovine serum albumin, 0.15 M NaCl, 0.03 M Tris-HCl (pH 8), 1 mM EDTA, and rinsed with 0.02% Ficoll, 0.02% polyvinylpyrrolidone, 0.02% bovine serum albumin, 0.75 M NaCl, 0.15 M Tris-HCl (pH 8), 5 mM EDTA and 0.5% SDS. The hybridization proceeded overnight at 68°C in a solution identical to the rinsing solution above using the ^{32}P-labelled probe which had been denatured at 100°C for 5 min prior to use. The hybridized filters were washed twice with 0.3 M NaCl, 0.06 M Tris-HCl (pH 8), 2 mM EDTA for 2 h at 68°C before air drying and autoradiography.

About 1350 clones, originating from the 800-900 DNA size class, were screened. Thirteen colonies, including pHFIF1, gave a positive result. These clones were designated G-HB101-pHFIF1 to 13 and their recombinant DNA molecules pHFIF1 to 13. One of the clones, pHFIF2, was hybridized with poly(A) mRNA containing F IF mRNA and assayed using DBM-cellulose paper (supra). Because the total IF-RNA activity was detected in the hybridized fraction and the unhybridized RNA did not contain any detectable activity, it is clear that clones identified by colony hybridization to a part of the pHFIF1 fragment also hybridize to F IF mRNA.

It is, of course, evident that this method of clone screening may be employed equally well on other clones containing DNA sequences arising from recombinant DNA technology, synthesis, natural sources or a combination thereof or clones containing DNA sequences related to any of the above DNA sequences by mutation, including single or multiple, base substitutions, insertions, inversions, or deletions. Therefore, such DNA sequences and their identification also fall within this invention. It is also to be understood that DNA sequences, which are not screened by the above DNA sequences, yet which as a result of their arrangement of nucleotides code for those polypeptides coded for by the above DNA sequences also fall within this invention.

CHARACTERIZATION OF THE F IF-RELATED RECOMBINANT PLASMIDS

The thirteen clones which were detected by colony hybridization were further characterized. A physical map of the inserts of these clones was constructed and the orientation of the inserts in the various clones was determined.

The physical maps of the plasmids were constructed by digestion with various restriction enzymes (New England Biolabs) in 10 mM Tris-HCl (pH 7.6), 7 mM $MgCl_2$ and 7 mM β-mercaptoethanol at 37°C by well-known procedures. The products of digestion were electrophoresed in 2.2% agarose or 6% polyacrylamide gels in 40 mM Tris-HOAc (pH 7.8), 20 mM EDTA. They were analyzed after visualization by staining with ethidiumbromide and compared with the detailed physical map of pBR322 (J.G. Sutcliffe, supra). Restriction maps of the different plasmids were constructed on the basis of these digestion patterns. These were refined by sequencing the DNA inserts in various of the plasmids, substantially by the procedure of A.M. Maxam and W. Gilbert, "A New Method For Sequencing DNA", Proc. Natl. Acad. Sci. USA, 74, pp. 560-564 (1977).

Plasmid DNA was prepared from various of the pHFIF1-13 in accordance with this invention by the method

of Kahn et al. (supra), employed previously herein to isolate the DNA from the sets of clones for screening. The isolated form I DNA was purified by neutral sucrose-gradient centrifugation as before and restricted by various restriction enzymes, essentially as recommended by the supplier (New England Biolabs).

Restricted DNA was dephosphorylated for 30 min at 65°C in the presence of 4 units bacterial alkaline phosphatase and 0.1% SDS. Following two phenol extractions and ethanol precipitation, the DNA was 5'-terminally labelled with γ-^{32}P-ATP (~3000 Ci/mmole) and polynucleotide kinase (P-L Biochemicals, Inc.).

For sequencing, labelled fragments were handled in two ways. Some were purified on a polyacrylamide gel prior to cleavage with a second restriction enzyme. Others were immediately cleaved with a second restriction enzyme. In both cases the desired fragments were separated on a polyacrylamide gel in Tris-borate-EDTA buffer. Figure 7 displays the various restriction fragments (the circles indicating the label and the arrow the direction of sequencing) and the sequencing strategy employed using pHFIF1, pHFIF3, pHFIF6 and pHFIF7.

The fragments were degraded according to the method of A.M. Maxam and W. Gilbert (supra). The products were fractionated on polyacrylamide gels of various concentrations and lengths in 50 mM Tris-borate, 1 mM EDTA (pH 8.3) at 900 V to 2000 V.

Each stretch of cDNA insert was sequenced from both strands and each restriction site which served as labelled terminus was sequenced using a fragment spanning it. The composite nucleotide sequence thus obtained for the coding strand of F IF DNA or gene and its corresponding amino acid sequence is depicted in Fig. 4. Because none of plasmids pHFIF1-13 contained the complete gene for fibroblast interferon, Fig. 4 results from a combination of the data from at least two such plasmids. In this regard, Fig. 5 displays the relationship of inserts pHFIF1,

pHFIF3, pHFIF6 and pHFIF7, the solid arrows or chevrons indicating the orientation of the various parts of the inserts.

Referring now to Fig. 4, the heteropolymeric part of the insert is flanked on one end by a segment rich in T's and by a string of A's (probably reflecting the polyA terminus of the mRNA). For reference the insert is numbered from first nucleotide of the composite insert to a nucleotide well into the untranslated section of the insert. An ATG initiation triplet at position 65-67 and a TGA termination triplet at position 626-628 define a reading frame uninterrupted by nonsense codons. Any other translatable sequence, i.e., in different reading frames, flanked by an ATG or a GTG and a termination signal is too short to code for a polypeptide of the expected size of F IF. Therefore, the region between nucleotides 65 and 625 most likely includes the nucleotide sequence for the composite gene that codes for F IF in accordance with this invention. This sequence does not exclude the possibility that modifications to the gene such as mutations, including single or multiple, base substitutions, deletions, insertions, or inversions may not have already occurred in the gene or may not be employed subsequently to modify its properties.

It should of course be understood that cloned cDNA from polyA RNA by the usual procedures (A. Efstratiadis et al., supra) lacks 5'-terminal nucleotides and may even contain artifactual sequences (R.I. Richards et al., "Molecular Cloning And Sequence Analysis Of Adult Chicken β-Globin cDNA", Nucleic Acids Research, 7, pp. 1137-46 (1979)). Therefore, it is not certain that the ATG located at nucleotides 65-67 is in fact the first ATG of authentic mRNA. However, for the purposes of the following description, it is assumed that the ATG at nucleotides 65-67 is the first ATG of authentic F IF DNA.

By comparing the polypeptide coded by this region of the insert with that sequence of 13 amino-terminal amino

acids of authentic human fibroblast interferon --MetSerTyr AsnLeuLeuGlyPheLeuGlnArgSerSer-- determined by Knight et al. (supra), it appears that the chosen reading frame is correct and that nucleotides 65-127 may code for a signal sequence which precedes the nucleotide sequence coding for the "mature" polypeptide. In addition, in eukaryotic mRNAs the first AUG triplet from the 5' terminus is usually the initiation site for protein synthesis (M. Kozak, "How Do Eukaryotic Ribosomes Select Initiation Regions In Messenger RNA?", Cell, 15, pp. 1109-25 (1978)). Here, the codon in the composite fragment corresponding to the first amino acid of fibroblast interferon is 22 codons from the first ATG. This again suggests that the DNA sequence coding for fibroblast interferon may be preceded by a sequence determining a signal polypeptide of 21 amino acids. The presumptive signal sequence contains a series of hydrophobic amino acids. An accumulation of hydrophobic residues is characteristic of signal sequences (cf., B.D. Davis and P.C. Tai, "The Mechanism Of Protein Secretion Across Membranes", Nature, 283, pp. 433-38 (1980)).

The nucleotide sequence apparently corresponding to "mature" F IF polypeptide comprises 498 nucleotides, which code for 166 amino acids. Assuming that there is no carboxyterminal processing, the molecular weight of the interferon polypeptide is 20085. The base composition of the coding sequence is 45% G+C. The codon usage within the interferon coding sequences is in reasonable agreement with that compiled for mammalian mRNAs in general (R. Grantham et al., "Coding Catalog Usage And The Genome Hypothesis", Nucleic Acids Research, 8, pp. 49-62 (1980)). Any deviations observed may be ascribed to the small numbers involved.

The structure of the polypeptide depicted in Fig. 4 for the composite fragment, of course, does not take into account any modifications to the polypeptide caused by its interaction with in vivo enzymes, e.g., glycosylation. Therefore, it must be understood that this structure may not be identical with F IF produced in vivo.

The comparison of the first 13 amino acids of authentic fibroblast interferon (Knight et al., supra) and the sequence deduced from the composite gene of Fig. 4 shows no differences. The amino acid compositions determined directly for authentic fibroblast interferon on the one hand and that deduced from the sequence of the composite gene of this invention on the other also show substantial similarties. Fig. 6 displays a comparison of these compositions.

Although none of the recombinant DNA molecules prepared in accordance with this invention contain the complete DNA sequence for fibroblast interferon, a combination of portions of the inserts of these recombinant DNA molecules to afford the complete F IF DNA gene sequence is within the skill of the art. For example, by reference to Fig. 5, it can readily be seen that the PstI-BglII fragment of pHFIF6 could be joined with the PstI-HaeII fragment of pHFIF7 or the EcoRI-PstI fragment of pHFIF6 could be joined with the PstI-HaeII fragment of pHFIF7 or the BglII-PstI fragment of pHFIF6 could be joined with the PstI-BglII fragment of clone 7 to form the composite F IF gene. The joining of these fragments could be done before or after insertion into a desired plasmid.

Micro-organisms and recombinant DNA molecules prepared by the processes described herein are exemplified by cultures deposited in the culture collection Deutsche Sammlung von Mikroorganism in Gottingen, West Germany on April 2, 1980, and identified as HFIF-A to C:

A: E. coli HB101 (G-pBR322(Pst)/HFIF3)

B: E. coli HB101 (G-pBR322(Pst)/HFIF6)

C: E. coli HB101 (G-pBR322(Pst)/HFIF7)

These cultures were assigned accession numbers DSM 1791-1793, respectively.

While we have herein before presented a number of embodiments of this invention, it is apparent that our basic construction can be altered to provide other embodiments which utilize the processes and compositions of this invention. Therefore, it will be appreciated that the

scope of this invention is to be defined by the claims appended hereto rather than the specific embodiments which have been presented herein before by way of example.

CLAIMS

1. A recombinant DNA molecule characterized by a structural gene selected from the group comprising the DNA inserts of G-pBR322(Pst)/HFIF3, G-pBR322(Pst)/HFIF6 or pBR322(Pst)/HFIF7, DNA sequences which hybridize to any of the foregoing DNA inserts, or DNA sequences, from whatever source obtained, including natural, synthetic, or semi-synthetic sources, related by mutation, including single or multiple, base substitutions, deletions, insertions and inversions, to any of the foregoing DNA sequences or inserts.

2. A recombinant DNA molecule characterized by a structural gene having the formula: ATGACCAACAAGTGTCTC CTCCAAATTGCTCTCCTGTTGTGCTTCTCCACTACAGCTCTTTCCATGAGCTACAAC TTGCTTGGATTCCTACAAAGAAGCAGCAATTTTCAGTGTCAGAAGCTCCTGTGGCAA TTGAATGGGAGGCTTGAATACTGCCTCAAGCACAGGATGAACTTTGACATCCCTGAG GAGATTAAGCAGCTGCAGCAGTTCCAGAAGGAGGACGCCGCATTGACCATCTATGAG ATGCTCCAGAACATCTTTGCTATTTTCAGACAAGATTCATCTAGCACTGGCTGGAAT GAGACTATTGTTGAGAACCTCCTGGCTAATGTCTATCATCAGATAAACCATCTGAAG ACAGTCCTGGAAGAAAAACTGGAGAAAGAAGATTTCACCAGGGGAAAACTCATGAGC AGTCTGCACCTGAAAAGATATTATGGGAGGATTCTGCATTACCTGAAGGCCAAGGAG TACAGTCACTGTGCCTGGACCATAGTCAGAGTGGAAATCCTAAGGAACTTTTACTTC ATTAACAGACTTACAGGTTACCTCCGAAAC.

3. A recombinant DNA molecule characterized by a structural gene having the formula: ATGAGCTACAACTTGCTT GGATTCCTACAAAGAAGCAGCAATTTTCAGTGTCAGAAGCTCCTGTGGCAATTGAAT GGGAGGCTTGAATACTGCCTCAAGCACAGGATGAACTTTGACATCCCTGAGGAGATT AAGCAGCTGCAGCAGTTCCAGAAGGAGGACGCCGCATTGACCATCTATGAGATGCTC CAGAACATCTTTGCTATTTTCAGACAAGATTCATCTAGCACTGGCTGGAATGAGACT ATTGTTGAGAACCTCCTGGCTAATGTCTATCATCAGATAAACCATCTGAAGACAGTC CTGGAAGAAAAACTGGAGAAAGAAGATTTCACCAGGGGAAAACTCATGAGCAGTCTG CACCTGAAAAGATATTATGGGAGGATTCTGCATTACCTGAAGGCCAAGGAGTACAGT CACTGTGCCTGGACCATAGTCAGAGTGGAAATCCTAAGGAACTTTTACTTCATTAAC AGACTTACAGGTTACCTCCGAAAC.

4. The recombinant DNA molecule according to claims 1 to 3, wherein the molecule comprises a cloning vehicle having a first and a second restriction endonu-

clease recognition site, said structural gene being inserted between the first and second restriction sites.

5. A recombinant DNA molecule according to claims 1 to 4, selected from the group comprising G-pBR322(Pst)/HFIF3, G-pBR322(Pst)/HFIF6 or G-pBR322(Pst)/HFIF7, molecules whose DNA inserts hybridize to the DNA inserts in any of the foregoing molecules, or molecules, from whatever source obtained, including natural, synthetic or semi-synthetic sources, related by mutation, including single or multiple, base substitutions, deletions, insertions and inversions to the DNA inserts from any of the foregoing molecules.

6. A recombinant DNA molecule characterized by a structural gene comprising a sequence of codons which codes for a polypeptide similar in amino acid sequence to those coded for by the codons of a structural gene selected from the group of genes of the formula: ATGACCAACAAGTGTCTC CTCCAAATTGCTCTCCTGTTGTGCTTCTCCACTACAGCTCTTTCCATGAGCTACAAC TTGCTTGGATTCCTACAAAGAAGCAGCAATTTTCAGTGTCAGAAGCTCCTGTGGCAA TTGAATGGGAGGCTTGAATACTGCCTCAAGCACAGGATGAACTTTGACATCCCTGAG GAGATTAAGCAGCTGCAGCAGTTCCAGAAGGAGGACGCCGCATTGACCATCTATGAG ATGCTCCAGAACATCTTTGCTATTTTCAGACAAGATTCATCTAGCACTGGCTGGAAT GAGACTATTGTTGAGAACCTCCTGGCTAATGTCTATCATCAGATAAACCATCTGAAG ACAGTCCTGGAAGAAAAACTGGAGAAAGAAGATTTCACCAGGGGAAAACTCATGAGC AGTCTGCACCTGAAAAGATATTATGGGAGGATTCTGCATTACCTGAAGGCCAAGGAG TACAGTCACTGTGCCTGGACCATAGTCAGAGTGGAAATCCTAAGGAACTTTTACTTC ATTAACAGACTTACAGGTTACCTCCGAAAC, ATGAGCTACAACTTGCTTGGATTCC TACAAAGAAGCAGCAATTTTCAGTGTCAGAAGCTCCTGTGGCAATTGAATGGGAGGC TTGAATACTGCCTCAAGCACAGGATGAACTTTGACATCCCTGAGGAGATTAAGCAGC TGCAGCAGTTCCAGAAGGAGGACGCCGCATTGACCATCTATGAGATGCTCCAGAACA TCTTTGCTATTTTCAGACAAGATTCATCTAGCACTGGCTGGAATGAGACTATTGTTG AGAACCTCCTGGCTAATGTCTATCATCAGATAAACCATCTGAAGACAGTCCTGGAAG AAAAACTGGAGAAAGAAGATTTCACCAGGGGAAAACTCATGAGCAGTCTGCACCTGA AAAGATATTATGGGAGGATTCTGCATTACCTGAAGGCCAAGGAGTACAGTCACTGTG CCTGGACCATAGTCAGAGTGGAAATCCTAAGGAACTTTTACTTCATTAACAGACTTA CAGGTTACCTCCGAAAC, DNA sequences which hybridize to any of the foregoing genes or DNA sequences, and DNA sequences

from whatever source obtained, including natural, synthetic or semi-synthetic sources, related by mutation, including single or multiple, base substitutions, deletions, insertions and inversions, to any of the foregoing genes or sequences..

7. A host transformed with at least one recombinant DNA molecule according to any of the preceding claims.

8. The transformed host of claim 7 characterized in that the host is selected from the group comprising strains of E. coli, Pseudomonas, Bacillus subtilis, Bacillus stearothermophilus, other bacilli, yeasts, other fungi, animal and plant hosts or human tissue cells.

9. The transformed host according to claims 7 to 8, characterized in that it comprises E. coli HB101 (G-pBR322(Pst)/HFIF3), E. coli HB101 (G-pBR322(Pst)/HFIF6), or E. coli HB101 (G-pBR322(Pst)/HFIF7).

10. A gene selected from the group comprising the DNA inserts of G-pBR322(Pst)/HFIF3, G-pBR322(Pst)/HFIF6 or G-pBR322(Pst)/HFIF7, DNA sequences which hybridize to any of the foregoing DNA inserts, or DNA sequences, from whatever source obtained, including natural, synthetic or semi-synthetic sources, related by mutation, including single or multiple, base substitutions, deletions, insertions and inversions to any of the foregoing DNA sequences or inserts.

11. A gene selected from the group of genes of the formula: ATGACCAACAAGTGTCTCCTCCAAATTGCTCTCCTGTTGTGCT
TCTCCACTACAGCTCTTTCCATGAGCTACAACTTGCTTGGATTCCTACAAAGAAGCA
GCAATTTTCAGTGTCAGAAGCTCCTGTGGCAATTGAATGGGAGGCTTGAATACTGCC
TCAAGCACAGGATGAACTTTGACATCCCTGAGGAGATTAAGCAGCTGCAGCAGTTCC
AGAAGGAGGACGCCGCATTGACCATCTATGAGATGCTCCAGAACATCTTTGCTATTT
TCAGACAAGATTCATCTAGCACTGGCTGGAATGAGACTATTGTTGAGAACCTCCTGG
CTAATGTCTATCATCAGATAAACCATCTGAAGACAGTCCTGGAAGAAAAACTGGAGA
AAGAAGATTTCACCAGGGGAAAACTCATGAGCAGTCTGCACCTGAAAAGATATTATG
GGAGGATTCTGCATTACCTGAAGGCCAAGGAGTACAGTCACTGTGCCTGGACCATAG
TCAGAGTGGAAATCCTAAGGAACTTTTACTTCATTAACAGACTTACAGGTTACCTGC
GAAAC, ATGAGCTACAACTTGCTTGGATTCCTACAAAGAAGCAGCAATTTTCAGTG

TCAGAAGCTCCTGTGGCAATTGAATGGGAGGCTTGAATACTGCCTCAAGCACAGGAT
GAACTTTGACATCCCTGAGGAGATTAAGCAGCTGCAGCAGTTCCAGAAGGAGGACGC
CGCATTGACCATCTATGAGATGCTCCAGAACATCTTTGCTATTTTCAGACAAGATTC
ATCTAGCACTGGCTGGAATGAGACTATTGTTGAGAACCTCCTGGCTAATGTCTATCA
TCAGATAAACCATCTGAAGACAGTCCTGGAAGAAAAACTGGAGAAAGAAGATTTCAC
CAGGGGAAAACTCATGAGCAGTCTGCACCTGAAAAGATATTATGGGAGGATTCTGCA
TTACCTGAAGGCCAAGGAGTACAGTCACTGTGCCTGGACCATAGTCAGAGTGGAAAT
CCTAAGGAACTTTTACTTCATTAACAGACTTACAGGTTACCTCCGAAAC, DNA sequences which hybridize to any of the foregoing genes, DNA sequences, from whatever source obtained, including natural, synthetic or semi-synthetic sources, related by mutation, including single or multiple, base substitutions, deletions, insertions and inversions to any of the foregoing genes or DNA sequences, or genes comprising a sequence of codons which codes for a polypeptide similar in amino acid sequence to those coded for by any of the foregoing DNA sequences or genes.

12. A screening process for DNA sequences characterized by the step of determining whether said DNA sequence hybridizes to at least one of the DNA inserts of G-pBR322(Pst)/HFIF3, G-pBR322(Pst)/HFIF6 or G-pBR322(Pst)/HFIF7, DNA sequences which hybridize to any of the foregoing DNA inserts and DNA sequences, from whatever source obtained, including natural, synthetic or semi-synthetic sources, related by mutation, including single or multiple, base substitutions, deletions, inversions and insertions to any of the foregoing DNA inserts or sequences.

13. A screening process for DNA sequences characterized by the step of determining whether said DNA sequences hybridizes to at least one of a gene selected from the group of ATGACCAACAAGTGTCTCCTCCAAATTGCTCTCCTGTTGTGCT
TCTCCACTACAGCTCTTTCCATGAGCTACAACTTGCTTGGATTCCTACAAAGAAGCA
GCAATTTTCAGTGTCAGAAGCTCCTGTGGCAATTGAATGGGAGGCTTGAATACTGCC
TCAAGCACAGGATGAACTTTGACATCCCTGAGGAGATTAAGCAGCTGCAGCAGTTCC
AGAAGGAGGACGCCGCATTGACCATCTATGAGATGCTCCAGAACATCTTTGCTATTT
TCAGACAAGATTCATCTAGCACTGGCTGGAATGAGACTATTGTTGAGAACCTCCTGG
CTAATGTCTATCATCAGATAAACCATCTGAAGACAGTCCTGGAAGAAAAACTGGAGA
AAGAAGATTTCACCAGGGGAAAACTCATGAGCAGTCTGCACCTGAAAAGATATTATG

GGAGGATTCTGCATTACCTGAAGGCCAAGGAGTACAGTCACTGTGCCTGGACCATAG
TCAGAGTGGAAATCCTAAGGAACTTTTACTTCATTAACAGACTTACAGGTTACCTCC
GAAAC, ATGAGCTACAACTTGCTTGGATTCCTACAAAGAAGCAGCAATTTTCAGTG
TCAGAAGCTCCTGTGGCAATTGAATGGGAGGCTTGAATACTGCCTCAAGCACAGGAT
GAACTTTGACATCCCTGAGGAGATTAAGCAGCTGCAGCAGTTCCAGAAGGAGGACGC
CGCATTGACCATCTATGAGATGCTCCAGAACATCTTTGCTATTTTCAGACAAGATTC
ATCTAGCACTGGCTGGAATGAGACTATTGTTGAGAACCTCCTGGCTAATGTCTATCA
TCAGATAAACCATCTGAAGACAGTCCTGGAAGAAAAACTGGAGAAAGAAGATTTCAC
CAGGGGAAAACTCATGAGCAGTCTGCACCTGAAAAGATATTATGGGAGGATTCTGCA
TTACCTGAAGGCCAAGGAGTACAGTCACTGTGCCTGGACCATAGTCAGAGTGGAAAT
CCTAAGGAACTTTTACTTCATTAACAGACTTACAGGTTACCTCCGAAAC, DNA se-
quences which hybridize to any of the foregoing genes, DNA
sequences, from whatever source obtained, including natural,
synthetic or semi-synthetic sources, related by mutation,
including single or multiple, base substitutions, deletions,
insertions and inversions to any of the foregoing genes or
DNA sequences.

14. The process of any of claims 12 to 13 char-
acterized in that the DNA sequence screened is selected
from the group comprising DNA sequences from natural
sources, synthetic DNA sequences, DNA sequences from
recombinant DNA molecules or DNA sequences, which are a
combination of the foregoing.

15. A method for producing a DNA sequence com-
prising the steps of preparing a recombinant DNA molecule
characterized by an inserted structural gene, said gene
being selected from the group comprising the DNA inserts
of G-pBR322(Pst)/HFIF3, G-pBR322(Pst)/HFIF6 or G-pBR322
(Pst)/HFIF7, DNA sequences which hybridize to any of the
foregoing DNA inserts, or DNA sequences, from whatever
source obtained, including natural, synthetic or semi-
synthetic sources related by mutation, including single
or multiple, base substitutions, deletions, insertions
and inversions, to any of the foregoing DNA sequences or
inserts, or DNA inserts which comprise a sequence of
codons which code for a polypeptide similar in amino
acid sequence to those polypeptides coded for by any of

the foregoing DNA inserts or sequences; transforming an appropriate host with said recombinant DNA molecule; culturing said host; and separating said DNA sequences.

16. A method for producing a DNA sequence comprising the steps of culturing a host transformed with at least one recombinant DNA molecule selected from the group comprising G-pBR322(Pst)/HFIF3, G-pBR322(Pst)/HFIF6, or G-pBR322(Pst)/HFIF7, molecules whose DNA inserts hybridize to the DNA inserts of any of the foregoing molecules, molecules whose DNA inserts, from whatever source obtained, including natural, synthetic or semi-synthetic sources, are related by mutation, including single or multiple, base substitutions, deletions, insertions and inversions to the DNA inserts of any of the foregoing molecules, or molecules whose DNA inserts comprise a sequence of codons which code for a polypeptide similar in amino acid sequence to those polypeptides coded for by the DNA inserts of any of the foregoing molecules.

17. A method for producing a DNA sequence comprising the steps of preparing a recombinant DNA molecule characterized by an inserted structural gene, said gene being selected from the group comprising

ATGACCAACAAGTGTCTCCTCCAAATTGCTCTCCTGTTGTGCT
TCTCCACTACAGCTCTTTCCATGAGCTACAACTTGCTTGGATTCCTACAAAGAAGCA
GCAATTTTCAGTGTCAGAAGCTCCTGTGGCAATTGAATGGGAGGCTTGAATACTGCC
TCAAGCACAGGATGAACTTTGACATCCCTGAGGAGATTAAGCAGCTGCAGCAGTTCC
AGAAGGAGGACGCCGCATTGACCATCTATGAGATGCTCCAGAACATCTTTGCTATTT
TCAGACAAGATTCATCTAGCACTGGCTGGAATGAGACTATTGTTGAGAACCTCCTGG
CTAATGTCTATCATCAGATAAACCATCTGAAGACAGTCCTGGAAGAAAAACTGGAGA
AACAAGATTTCACCAGGGGAAAACTCATGAGCAGTCTGCACCTGAAAAGATATTATG
GGAGGATTCTGCATTACCTGAAGGCCAAGGAGTACAGTCACTGTGCCTGGACCATAG
TCAGAGTGGAAATCCTAAGCAACTTTTACTTCATTAACAGACTTACAGGTTACCTCC
GAAAC, ATGAGCTACAACTTGCTTGGATTCCTACAAAGAAGCAGCAATTTTCAGTG
TCAGAAGCTCCTGTGGCAATTGAATGGGAGGCTTGAATACTGCCTCAAGCACAGGAT
GAACTTTGACATCCCTGAGGAGATTAAGCAGCTGCAGCAGTTCCAGAAGGAGGACGC
CGCATTGACCATCTATGAGATGCTCCAGAACATCTTTGCTATTTTCAGACAAGATTC
ATCTAGCACTGGCTGGAATGAGACTATTGTTGAGAACCTCCTGGCTAATGTCTATCA
TCAGATAAACCATCTGAAGACAGTCCTGGAAGAAAAACTGGAGAAAGAAGATTTCAC

CAGGGGAAAACTCATGAGCAGTCTGCACCTGAAAAGATATTATGGGAGGATTCTGCA
TTACCTGAAGGCCAAGGAGTACAGTCACTGTGCCTGGACCATAGTCAGAGTGGAAAT
CCTAAGGAACTTTTACTTCATTAACAGACTTACAGGTTACCTCCGAAAC, DNA se-
quences which hybridize to any of the foregoing genes, DNA
sequences, from whatever source obtained, including natural,
synthetic or semi-synthetic sources, related by mutation,
including single or multiple, base substitutions, deletions,
insertions and inversions to any of the foregoing genes or
DNA sequences, or DNA sequences comprising a sequence of
codons which codes for a polypeptide similar in amino acid
sequence to those polypeptides coded for by the codons of
any of the foregoing genes or DNA sequences; transforming
an appropriate host with said recombinant DNA molecule;
culturing said host and separating said DNA sequence.

18. A method for producing a DNA sequence com-
prising the step of culturing a host transformed with at
least one recombinant DNA molecule selected from the group
comprising molecules whose DNA inserts are selected from the
group of genes comprising

ATGACCAACAAGTGTCTCCTCCAAATTGCTCTCCTGTTGTGCT
TCTCCACTACAGGTCTTTCCATGAGCTACAACTTGCTTGGATTCCTACAAAGAAGCA
GCAATTTTCAGTGTCAGAAGCTCCTGTGGCAATTGAATGGGAGGCTTGAATACTGCC
TCAAGCACAGGATGAACTTTGACATCCCTGAGGAGATTAAGCAGCTGCAGCAGTTCC
AGAAGGAGGACGCCGCATTGACCATCTATGAGATGCTCCAGAACATCTTTGCTATTT
TCAGACAAGATTCATCTAGCACTGGCTGGAATGAGACTATTGTTGAGAACCTCCTGG
CTAATGTCTATCATCAGATAAACCATCTGAAGACAGTCCTGGAAGAAAAACTGGAGA
AAGAAGATTTCACCAGGGGAAAACTCATGAGCAGTCTGCACCTGAAAAGATATTATG
GGAGGATTCTGCATTACCTGAAGGCCAAGGAGTACAGTCACTGTGCCTGGACCATAG
TCAGAGTGGAAATCCTAAGGAACTTTTACTTCATTAACAGACTTACAGGTTACCTCC
GAAAC, ATGAGCTACAACTTGCTTGGATTCCTACAAAGAAGCAGCAATTTTCAGTG
TCAGAAGCTCCTGTGGCAATTGAATGGGAGGCTTGAATACTGCCTCAAGCACAGGAT
GAACTTTGACATCCCTGAGGAGATTAAGCAGCTGCAGCAGTTCCAGAAGGAGGACGC
CGCATTGACCATCTATGAGATGCTCCAGAACATCTTTGCTATTTTCAGACAAGATTC
ATCTAGCACTGGCTGGAATGAGACTATTGTTGAGAACCTCCTGGCTAATGTCTATCA
TCAGATAAACCATCTGAAGACAGTCCTGGAAGAAAAACTGGAGAAAGAAGATTTCAC
CAGGGGAAAACTCATGAGCAGTCTGCACCTGAAAAGATATTATGCGAGGATTCTGCA
TTACCTGAAGGCCAAGGAGTACAGTCACTGTGCCTGGACCATAGTCAGAGTGGAAAT
CCTAAGGAACTTTTACTTCATTAACAGACTTACAGGTTACCTCCGAAAC, DNA se-

quences which hybridize to any of the foregoing genes, DNA sequences, from whatever source obtained, including natural, synthetic or semi-synthetic sources, related by mutation, including single or multiple, base substitutions, deletions, insertions and inversions to any of the foregoing genes or DNA sequences or DNA sequences which code for a polypeptide similar in amino acid sequence to those polypeptides coded for by any of the foregoing genes or DNA sequences.

19. The method of any of claims 15 to 18 characterized in that the host is selected from the group comprising strains of E. coli, Pseudomonas, Bacillus subtilis, Bacillus stearothermophilus, other bacilli, yeasts, other fungi, animal and plant hosts, or human tissue cells.

RECOMBINANT DNA MOLECULES AND THEIR USE IN PRODUCING STRUCTURAL GENES FOR HUMAN FIBROBLAST INTERFERON

ABSTRACT

Recombinant DNA molecules and hosts transformed with them which contain and produce structural genes for human fibroblast interferon and methods of making and using these molecules, host and genes. The recombinant DNA molecules are characterized by structural genes for human fibroblast interferon and fragments thereof.

FIG. 4

Amino Acid Composition of Human Fibroblast Interferon

Amino Acid	Composition			
	from direct analysis by Tan et al.	from direct analysis by Knight et al.	deduced from nucleotide sequence	
Asp	20.6	18.9	5	17
Asn			12	
Thr	8.0	6.8	7	
Ser	11.7	10.5	9	
Glu	27.5	27.0	13	24
Gln			11	
Pro	4.4	2.7	1	
Gly	5.4	7.8	6	
Ala	9.3	10.0	6	
Cys	N.D.	1.7	3	
Val	7.9	6.0	5	
Met	trace	2.9	4	
Ile	10.0	9.0	11	
Leu	26.9	20.4	24	
Tyr	3.2	7.5	10	
Phe	7.7	9.4	9	
His	4.6	4.9	5	
Lys	12.3	11.6	11	
Arg	8.6	10.9	11	
Trp	0.0	1.0	3	
TOTAL	168	169	166	

Fig. 6

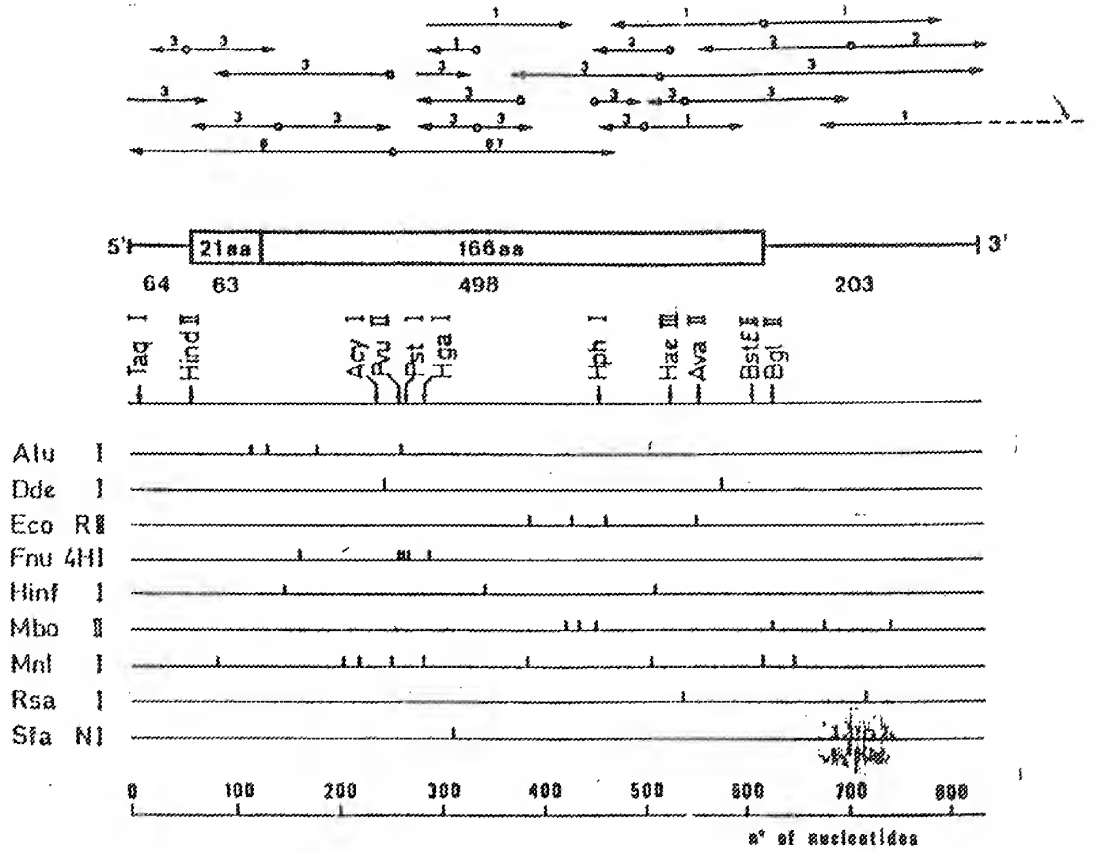
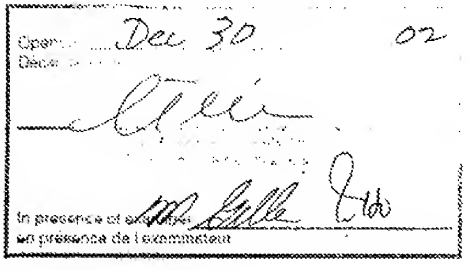
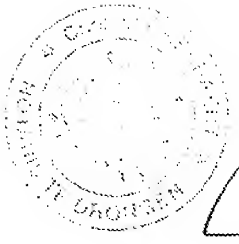


FIG. 7

BIOGEN GB PRIORITY APPLICATION

8011306





This is EXHIBIT FIERS-19
to
the Affidavit of Walter C. Fiers
sworn before me
this 13 th day of November, 2001

Commissioner for Oath or Notary Public

THE PATENT OFFICE,
25 SOUTHAMPTON BUILDINGS,
LONDON.

I, the undersigned, being an officer duly authorised in accordance with Section 62(3) of the Patents and Designs Act, 1907, to sign and issue certificates on behalf of the Comptroller-General, hereby certify that annexed hereto is a true copy of documents as originally filed in connection with the Patent application identified therein.

WITNESS my hand this
30 day of MARCH 1961

COC.

PATENTS ACT 1977 3 APRIL 1980

PATENTS FORM NO. 1/77
(Rules 6, 15, 18)

PATENT OFFICE

The Comptroller
The Patent Office
25 Southampton Buildings
London WC2A 1AY

1980
11306

8011306

REQUEST FOR THE GRANT OF A PATENT

THE GRANT OF A PATENT IS REQUESTED BY THE UNDERSIGNED ON THE BASIS OF THE PRESENT APPLICATION

I Applicant's or Agent's Reference *(Please insert if available)* EP/LJ

II Title of Invention Recombinant DNA Molecules and their use in producing structural genes for human fibroblast interferon

III Applicant or Applicants *(See note 2)*

Name (First or only applicant) BIOGEN N.V.

Address 24 Handelskade, Willemstad, Curacao, Netherlands Antilles

Nationality Netherlands Antilles Company

Name (of second applicant, if more than one)

Address

Nationality

IV Inventor *(See note 3)*

~~(a) The applicant(s) is the sole inventor(s)~~
or
(b) A statement on Patents Form No. 7/77 *will* be furnished

V Authorisation of Agent *(See note 4)* MEWBURN ELLIS & CO.

VI Address for Service *(See note 5)* 70/72 Chancery Lane
London W.C.2.

VII Declaration of Priority *(See note 6)*

Country	Filing date	File number

VIII The Application claims an earlier date under Section 8(3), 12(6), 15(4) or 37(4) *(See note 7)*

Earlier application or patent number ..... and filing date .....

C3554

IX Check List *(To be filled in by applicant or agent)*

A The application contains the following number of sheet(s)			B The application as filed is accompanied by:	
1	Request	1 Sheet(s)	1	Priority document 0
2	Description	42 Sheet(s)	2	Translation of priority document 0
3	Claim(s)	8 Sheet(s)	3	Request for Search 0
4	Drawing(s)	7 informal Sheet(s)	4	Statement of Inventorship and Right to Apply 0
5	Abstract	1 Sheet(s)	5	Separate Authorisation of To follow

X It is suggested that Figure No 1 of the drawings (if any) should accompany the abstract when published

XI Signature *(See note 8)*

Melvin Mitchell

NOTES:

1 This form, when completed, should be brought or sent to the Patent Office together with the prescribed fee and two copies of the description of the invention.

2 The name, address and nationality of each applicant are to be stated in the spaces provided at (i). Names of natural persons should be indicated in full. Bodies corporate should be designated by their corporate name. If there are more than two applicants the information concerning the third (and further) applicants should be given on a separate sheet.

3 Where the applicant or applicants is/are the sole inventor or the joint inventors, the declaration (a) to that effect at IV should be completed and the alternative statement (b) deleted. If however this is not the case the declaration (a) should be struck out and a statement will then be required to be filed upon Patents Form No. 7/77.

4 If the applicant wishes to appoint an agent, his name and address of his place of business shall be indicated in the spaces available at V and VI; such indication will be considered to be an authorisation for the agent to prosecute the application up to grant of a patent and to service any patent so granted.

5 If no authorised agent is appointed an address for service in the United Kingdom to which all documents and notices may be sent must be stated at VI. It is recommended that a telephone number be provided if available.

6 The declaration of priority at VII should state the date of the previous filing and the country in which it was made and indicate the file number, if available.

7 When an application is made by virtue of section 8(3), 12(6), 15(4) or 37(4) the appropriate section should be identified at VIII and the number of the earlier application or any patent granted thereon identified.

8 An agent may sign only when previously authorised. An express authorisation signed by the applicant(s) must be received by the Patent Office before the expiry of 3 months from the filing date.

9 Attention of applicants is drawn to the desirability of avoiding publication of inventions relating to any article, material or device intended or adapted for use in war (Official Secrets Acts, 1911 and 1920). In addition after an application for a patent has been filed at the Patent Office the comptroller will consider whether publication or communication of the invention should be prohibited or restricted under section 22 of the Act and will inform the applicant if such prohibition is necessary.

10 Applicants resident in the United Kingdom are also reminded that, under the provisions of section 23 applications may not be filed abroad without written permission or unless an application has been filed not less than six weeks previously in the United Kingdom for a patent for the same invention and no direction prohibiting publication or communication has been given or any such direction has been received.

Walter Charles FIERS
Benkendreef 3
B-9120 Destelbergen,
Belgium.

NOTES

1 The name(s) and address(es) of the inventor(s) are to be inserted in the spaces provided alongside.

2 Where more than 3 inventors are to be named, the names of the 4th and any further inventors should be given on the reverse side of an additional blank copy of Patents Form No.7/77 and attached to this form.

RECOMBINANT DNA MOLECULES AND THEIR USE IN PRODUCING STRUCTURAL GENES FOR HUMAN FIBROBLAST INTERFERON

TECHNICAL FIELD OF INVENTION

This invention relates to recombinant DNA molecules and their use in producing structural genes for human fibroblast interferon. The recombinant DNA molecules disclosed herein are characterized by DNA sequences that code for polypeptides whose amino acid sequence and composition are substantially consistent with human fibroblast interferon.

BACKGROUND ART

Two classes of interferons ("IF") are known to exist. Interferons of Class I are small, acid stable (glyco)-proteins that render cells resistant to viral infection (A. Isaacs and J. Lindenmann, "Virus Interference I. The Interferon", Proc. Royal Soc. Ser. B., 147, pp. 258-67 (1957) and W. E. Stewart, II, The Interferon System, Springer-Verlag (1979) (hereinafter "The Interferon System")). Class II IFs are acid labile. At present, they are poorly characterized. Although to some extent cell specific (The Interferon System, pp. 135-45), IFs are not virus specific. Instead, IFs protect cells against a wide spectrum of viruses.

Two antigenically distinct species of Class I human interferon ("HIF") are known to exhibit IF activity. One IF species, fibroblast interferon ("F IF"), is produced upon appropriate induction in diploid fibroblast cells. Another IF species, leukocyte interferon ("Le IF") is produced together with minor amounts of F IF upon appropriate induction in human leukocyte and lymphoblastoid cells. Both are heterogeneous in regard to size, presumably because of the carbohydrate moiety. F IF has been extensively purified and characterized

 DUPLICATE

(E. Knight, Jr., "Interferon: Purification And Initial Characterization From Human Diploid Cells", Proc. Natl. Acad. Sci. USA, 73, pp. 520-23 (1976)). It is a glycoprotein of about 20,000 molecular weight (M. Wiranowska-Stewart, et al., "Contributions Of Carbohydrate Moieties To The Physical And Biological Properties Of Human Leukocyte, Lymphoblastoid And Fibroblast Interferons", Abst. Ann. Meeting Amer. Soc. Microbiol., p. 246 (1978)). Its amino acid composition has been determined (E. Knight, Jr., et al., "Human Fibroblast Interferon: Amino Acid Analysis And Amino-Terminal Amino Acid Sequence", Science, 207, pp. 525-26 (1980)). Elucidation of its amino acid sequence is in progress. To date, the amino acid sequence of the NH_2 terminus of the mature protein has been reported for the first 13 amino acid residues: Met-Ser-Tyr-Asn-Leu-Leu-Gly-Phe-Leu-Gln-Arg-Ser-Ser... (E. Knight, Jr., et al., supra). Two distinct genes, one located on chromosome 2, the other on chromosome 5, have been reported to code for F IF (D. L. Slate and F. H. Ruddle, "Fibroblast Interferon In Man Is Coded By Two Loci On Separate Chromosomes", Cell, 16, pp. 171-80 (1979)). Other studies, however, indicate that the gene for F IF is located on chromosome 9 (A. Medger, et al., "Involvement Of A Gene On Chromosome 9 In Human Fibroblast Interferon Production", Nature, 280, pp. 493-95 (1979)).

Le IF has likewise been purified and characterized. Two components have been described, one of 21000 to 22000 and the other of 15000 to 18000 molecular weight (K. C. Zoon, et al., "Purification And Partial Characterization Of Human Lymphoblastoid Interferon", Proc. Natl. Acad. Sci. USA, 76, pp. 5601-605 (1979)). A portion of the amino acid sequence of Le IF has also been determined, i.e., 20 amino acids from the amino terminus of the mature protein (K. C. Zoon, et al., "Amino-Terminal Sequence Of The Major Component Of Human

Lymphoblastoid Interferon", Science, 207, pp. 527-28 (1980)).

A comparison of the initial amino acid sequence of F IF and Le IF reveals no detectable homology within the first 13 amino acids. The total amino acid compositions of the two species are also distinct. In addition, degradation of the sugar residues of the two species by periodate indicates that the carbohydrate structure of the two IFs is different (M. Wiranowska-Stewart, et al., supra).

The two species of HIF have a number of different properties. For example, anti-human Le IF antibodies are less efficient against F IF and anti-sera to human F IF have no activity against human Le IF (The Interferon System, p. 151). Le IF displays a high degree of activity in cell cultures of bovine, feline or porcine origin whereas F IF is hardly active in those cells but has been reported to be active in rat cells (P. Duc-Goiran, et al., "Studies On Virus-Induced Interferons Produced By The Human Aminotic Membrane And White Blood Cells", Arch. Ges. Virus Forsch., 34, pp. 232-43 (1971)). In addition, the two IFs result from different mRNA species (and therefore from presumably different structural genes) that code for polypeptides of different primary sequence (R. L. Cavelieri, et al., "Synthesis of Human Interferon By Xenopus laevis Oocytes: Two Structural Genes For Interferon In Human Cells", Proc. Natl. Acad. Sci. USA, 74, pp. 3287-91 (1977)).

Although both Le and F IFs occur in a glycosylated form, removal of the carbohydrate moiety (P. J. Bridgen, et al., "Human Lymphoblastoid Interferon", J. Biol. Chem., 252, pp. 6585-87 (1977)) or synthesis of IF in the presence of inhibitors which preclude glycosylation (W. E. Stewart, II, et al., "Effect of Glycosylation Inhibitors On The Production And Properties Of Human Leukocyte Interferon", Virology, 97, pp. 473-76

(1979); J. Fujisawa, et al., "Nonglycosylated Mouse L Cell Interferon Produced By The Action Of Tunicamycin", J. Biol. Chem., 253, pp. 8677-79 (1978); E. A. Havell, et al., "Altered Molecular Species Of Human Interferon Produced In The Presence Of Inhibitors of Glycosylation", J. Biol. Chem., 252, pp. 4425-27 (1977); The Interferon System, p. 181) yields a smaller form of IF which still retains most or all of its IF activity.

Both F IF and Le IF may, like many human proteins, be polymorphic. Therefore, cells of particular individuals may produce IF species within each of the more general F IF and Le IF classes which are physiologically similar but structurally slightly different than the prototype of the class of which it is a part. Therefore, while the protein structure of the F IF or Le IF may be generally well-defined, particular individuals may produce IFs that are slight variations thereof.

IF is usually not detectable in normal or healthy cells (The Interferon System, pp. 55-57). Instead, the protein is produced as a result of the cell's exposure to an IF inducer. IF inducers are usually viruses but may also be non-viral in character, such as natural or synthetic double-stranded RNA, intracellular microbes, microbial products and various chemical agents. Numerous attempts have been made to take advantage of these non-viral inducers to render human cells resistant to viral infection (S. Baron and F. Dianzani (eds.), Texas Reports On Biology And Medicine, 35 ("Texas Reports"), pp. 528-40 (1977)). These attempts have not been very successful. Instead, use of exogenous IF itself is now preferred.

As an antiviral agent, HIF has been used to treat the following: respiratory infections (Texas Reports, pp. 486-96); herpes simplex keratitis (Texas Reports, pp. 497-500; R. Sundmacher, "Exogenous Interferon in Eye Diseases", International Virology IV, The

Hague, Abstract nr. W2/11, p. 99 (1978)); acute hemorrhagic conjunctivitis (Texas Reports, pp. 501-10); adenovirus keratoconjunctivitis (A. Romano, et al., ISH Memo I-A8131 (October, 1979)); varicella zoster (Texas Reports, pp. 511-15); cytomegalovirus infection (Texas Reports, pp. 523-27); and hepatitis B (Texas Reports, pp. 516-22). See also The Interferon System, pp. 307-19. In these treatments F IF and Le IF may display different dose/response curves. However, large-scale use of IF as an antiviral agent requires larger amounts of HIF than heretofore have been available.

IF has other effects in addition to its antiviral action. For example, it antagonizes the effect of colony stimulating factor, inhibits the growth of hemopoietic colony-forming cells and interferes with the normal differentiation of granulocyte and macrophage precursors (Texas Reports, pp. 343-49). It also inhibits erythroid differentiation in DMSO-treated Friend leukemia cells (Texas Reports, pp. 423-28). Some cell lines may be considerably more sensitive to F IF than to Le IF in these regards (S. Einhorn and H. Strander, "Is Interferon Tissue- Specific? - Effect Of Human Leukocyte And Fibroblast Interferons On The Growth Of Lymphoblastoid And Osteosarcoma Cell Lines", J. Gen. Virol., 35, pp. 573-77 (1977); T. Kuwata, et al., "Comparison Of The Suppression Of Cell And Virus Growth In Transformed Human Cells By Leukocyte And Fibroblast Interferon", J. Gen. Virol., 43, pp. 435-39 (1979)).

IF may also play a role in regulation of the immune response. For example, depending upon the dose and time of application in relation to antigen, IF can be both immunopotentiating and immunosuppressive in vivo and in vitro (Texas Reports, pp. 357-69). In addition, specifically sensitized lymphocytes have been observed to produce IF after contact with antigen. Such antigen-induced IF could therefore be a regulator of the immune

response, affecting both circulating antigen levels and expression of cellular immunity (Texas Reports pp. 370-74). IF is also known to enhance the activity of killer lymphocytes and antibody-dependent cell-mediated cytotoxicity (R. R. Herberman, et al., "Augmentation By Interferon Of Human Natural And Antibody-Dependent Cell-Mediated Cytotoxicity", Nature, 277, pp. 221-23 (1979); P. Beverley and D. Knight, "Killing Comes Naturally", Nature, 278, pp. 119-20 (1979). Texas Reports, pp. 375-80; J. R. Huddlestone, et al., "Induction And Kinetics Of Natural Killer Cells in Humans Following Interferon Therapy", Nature, 282, pp. 417-19 (1979); S. Einhorn, et al., "Interferon And Spontaneous Cytotoxicity In Man. II. Studies In Patients Receiving Exogenous Leukocyte Interferon", Acta Med. Scand., 204, pp. 477-83 (1978)). Both may be directly or indirectly involved in the immunological attack on tumor cells.

Therefore, in addition to its use as a human antiviral agent, HIF has potential application in anti-tumor and anticancer therapy (The Interferon System, pp. 319-21 and 394-99). It is now known that IFs affect the growth of many classes of tumors in many animals (The Interferon System, pp. 292-304). They, like other anti-tumor agents, seem most effective when directed against small tumors. The antitumor effects of animal IF are dependent on dosage and time but have been demonstrated at concentrations below toxic levels. Accordingly, numerous investigations and clinical trials have been and continue to be conducted into the antitumor and anticancer properties of HIFs. These include treatment of several malignant diseases such as osteosarcoma, acute myeloid leukemia, multiple myeloma and Hodgkin's disease (Texas Reports, pp. 429-35). In addition, F IF has recently been shown to cause local tumor regression when injected into subcutaneous tumoral nodules in melanoma and breast carcinoma-affected patients (T. Nemoto,

et al., "Human Interferons And Intralesional Therapy Of Melanoma And Breast Carcinoma", Amer. Assoc. For Cancer Research, Abs nr. 993, p. 246 (1979)). Significantly some cell lines which resist the anticellular effects of Le IF remain sensitive to F IF. This differential effect suggests that F IF may be usefully employed against certain classes of resistant tumor cells which appear under selective pressure in patients treated with high doses of Le IF (T. Kuwata, et al., supra; A. A. Creasy, et al., "The Role of G_0-G_1 Arrest In The Inhibition Of Tumor Cell Growth By Interferon", Abstracts, Conference On Regulatory Functions Of Interferons, N.Y. Acad. Sci., nr. 17 (October 23-26, 1979)). Although the results of these clinical tests are encouraging, the antitumor and anticancer applications of HIF have been severely hampered by lack of an adequate supply of purified HIF.

At the biochemical level IFs induce the formation of at least 3 proteins, a protein kinase (B. Lebleu, et al., "Interferon, Double-Stranded RNA And Protein Phosphorylation", Proc. Natl. Acad. Sci. USA, 73, pp. 3107-11 (1976); A. G. Hovanessian and I. M. Kerr, "The (2'-5') Oligoadenylate (ppp A2'-5A2'-5'A) Synthetase And Protein Kinase(s) From Interferon-Treated Cells", Eur. J. Biochem., 93, pp. 515-26 (1979)), a (2'-5')oligo(A) polymerase (A. G. Hovanessian, et al., "Synthesis Of Low-Molecular Weight Inhibitor Of Protein Synthesis With Enzyme From Interferon-Treated Cells", Nature, 268, pp. 537-39 (1977); A. G. Hovanessian and I. M. Kerr, Eur. J. Biochem, supra) and a phosphodiesterase (A. Schmidt, et al., "An Interferon-Induced Phosphodiesterase Degrading (2'-5')oligoisoadenylate And The C-C-A Terminus Of tRNA", Proc. Natl. Acad. Sci. USA, 76, pp. 4788-92 (1979)). Both F IF and Le IF appear to trigger similar enzymatic pathways (C. Baglioni, "Interferon-Induced Enzymatic Activities And Their Role In The Antiviral

State", Cell, 17, pp. 255-64 (1979)) and both may share a common active core because they both recognize a chromosome 21-coded cell receptor (M. Wiranowska-Stewart, "The Role Of Human Chromosome 21 In Sensitivity To Interferons", J. Gen. Virol., 37, pp. 629-34 (1977)). The appearance of one or more of these enzymes in cells treated with IF should allow a further characterization of proteins with IF-like activity.

Today, F IF is produced by human cell lines grown in tissue culture. It is a low yield, expensive process. One large producer makes only 40-50 x 10^8 units of crude F IF per year (V. G. Edy, et al., "Human Interferon: Large Scale Production In Embryo Fibroblast Cultures", in Human Interferon (W. R. Stinebring and P. J. Chapple, eds.), Plenum Publishing Corp., pp. 55-60 (1978)). On purification by adsorption to controlled pore glass beads, F IF of specific activity of about 10^6 units/mg may be recovered in 50% yield from the crude cell extracts (A. Billiau, et al., "Human Fibroblast Interferon for Clinical Trials: Production, Partial Purification And Characterization", Antimicrobial Agents And Chemotherapy, pp. 49-55 (1979)). Further purification to a specific activity of about 10^9 units/mg is accomplished by zinc chelate affinity chromatography in about 100% yield (A. Billiau, et al., "Production, Purification And Properties Of Human Fibroblast Interferon", Abstracts, Conference On Regulatory Functions Of Interferons, N.Y. Acad. Sci., nr 29 (October 23-26, 1979)). Because the specific activity of F IF is so high, the amount of F IF required for commercial applications is low. For example, 100 g of pure IF would provide between 3 and 30 million doses.

Recent advances in molecular biology have made it possible to introduce the DNA coding for specific non-bacterial eukaryotic proteins into bacterial cells. In general, with DNA other than that prepared via chemical

synthesis, the construction of such recombinant DNA molecules comprises the steps of producing a single-stranded DNA copy (cDNA) of a purified messenger RNA (mRNA) template for the desired protein; converting the cDNA to double-stranded DNA; linking the DNA to an appropriate site in an appropriate cloning vehicle to form a recombinant DNA molecule and transforming an appropriate host with that recombinant DNA molecule. Such transformation may permit the host to produce the desired protein. Several non-bacterial genes and proteins have been obtained in E. coli using recombinant DNA technology. These include, for example, Le IF (C. Weissmann, et al., Seminar, Massachusetts Institute of Technology, January 16, 1980). In addition, recombinant DNA technology has been employed to produce a plasmid said to contain a gene sequence coding for F IF (T. Taniguchi, et al., "Construction And Identification Of A Bacterial Plasmid Containing The Human Fibroblast Interferon Gene Sequence", Proc. Japan Acad. Ser. B, 55, pp. 464-69 (1979)).

However, in neither of the foregoing has the actual gene sequence of F IF been described and in neither has that sequence been compared to the initial amino acid sequence or amino acid composition of authentic F IF. The former work is directed only to Le IF, a distinct chemical, biological and immunological Class I interferon from F IF (cf. supra). The latter report is based solely on hybridization data. These data do not enable one to determine if the selected clone contains the complete or actual gene sequence coding for F IF or if the cloned gene sequence will be able to express F IF in bacteria. Hybridization only establishes that a particular DNA insert is to some extent homologous with and complementary to a mRNA component of the poly(A)RNA that induces interferon activity when injected into oocytes. Moreover, the extent of any homology is dependent

on the hybridization conditions chosen for the screening process. Therefore, hybridization to a mRNA component of poly(A) RNA alone does not demonstrate that the selected DNA sequence is a sequence which codes for F IF or a polypeptide which displays the immunological or biological activity of F IF.

At a seminar in Zurich on February 25, 1980, Taniguchi stated that he had determined the nucleotide sequence for his hybridizing clone. He also stated that the first 13 amino acids coded for by that sequence were identical to that determined by Knight, et al., supra, for authentic F IF. Taniguchi did not disclose the full nucleotide sequence for his clone or compare its amino acid composition with that determined for authentic F IF. Nor is this invention addressed as is the apparent suggestion of Research Disclosure No. 18309, pp. 361-62 (1979) to prepare pure or substantially pure IF mRNA before attempting to clone the HIF gene.

DISCLOSURE OF THE INVENTION

The present invention avoids the uncertainties referred to by providing the identification and a source of a structural gene whose nucleotide sequence is substantially consistent with the known amino acid composition and sequence of authentic F IF.

By virtue of this invention, it is therefore possible to obtain a structural gene that codes for a polypeptide whose amino acid sequence and composition is substantially consistent with authentic F IF. Replication of these genes in appropriate recombinant DNA molecule-host combinations permits the production of large quantities of these genes. These genes are useful, either as produced in the host or after appropriate derivatization or modification, in compositions and methods for detecting and improving the production of

these products themselves and in selecting other genes related thereto.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 is a schematic outline of one embodiment of a process of this invention for preparing a mixture of recombinant DNA molecules, some of which are characterized by inserted DNA sequences that characterize this invention.

Figure 2 is a schematic outline of the initial clone screening process of this invention.

Figure 3 is a schematic outline of one embodiment of a clone screening process using DNA sequences prepared in accordance with the invention.

Figure 4 displays the nucleotide sequence of a composite DNA insert to a recombinant DNA molecule of this invention. The sequence is numbered from the beginning of the insert well into the untranslated area of the insert. Nucleotides 65-127 represent a signal sequence and nucleotides 128-625 represent the "mature" fibroblast interferon. The amino acid sequences of the signal polypeptide are depicted above their respective nucleotide sequences; the amino acids of the signal polypeptide being numbered from -21 to -1 and its other mature interferon from 1 to 166. ~~Various restriction endonuclease cuts for this gene are also depicted in Figure 4.~~

Figure 5 displays the orientation and restriction maps of several plasmids in accordance with this invention.

Figure 6 is a comparison of the amino acid composition of human fibroblast interferon as determined in accordance with this invention and that determined from authentic fibroblast interferon.

Figure 7 displays a restriction map of the FIF gene of this invention and the sequencing strategy used in sequencing pHFIF3, pHFIF6 and pHFIF7.

BEST MODE OF CARRYING OUT THE INVENTION

In order that the invention herein described may be more fully understood, the following detailed description is set forth.

In the description the following terms are employed:

Nucleotide--A monomeric unit of DNA or RNA consisting of a sugar moiety (pentose), a phosphate, and a nitrogenous heterocyclic base. The base is linked to the sugar moiety via the glycosidic carbon (1' carbon of the pentose) and that combination of base and sugar is called a nucleoside. The base characterizes the nucleotide. The four DNA bases are adenine ("A"), guanine ("G"), cytosine ("C"), and thymine ("T"). The four bases are A, G, C and uracil ("U").

DNA Sequence--A linear array of nucleotides connected one to the other by phosphodiester bonds between the 3' and 5' carbons of adjacent pentoses.

Codon--A DNA sequence of three nucleotides (a triplet) which encodes through mRNA an amino acid, a translation start signal or a translation termination signal. For example, the nucleotide triplets TTA, TTG, CTT, CTC, CTA and CTG encode for the amino acid leucine ("Leu"). TAG, TAA and TGA are translation stop signals and ATG is a translation start signal.

Reading Frame--The grouping of codons during translation of mRNA into amino acid sequences. During translation the proper reading frame must be maintained. For example, the sequence GCTGCTTGTAAG may be translated in three reading frames or phases, each of which affords a different amino acid sequence:

GCT GCT TGT AAG--Ala-Gly-Cys-Lys
G CTG CTT GTA AG--Leu-Val-Val
GC TGC TTG TAA G--Trp-Leu-(STOP)

Polypeptide--A linear array of amino acids connected one to the other by peptide bonds between the α-amino and carboxy groups of adjacent amino acids.

Genome--The entire DNA of a cell or a virus. It includes inter alia the structural genes coding for the polypeptides of the substance, as well as operator, promoter and ribosome binding and interaction sequences, including sequences such as the Shine-Dalgarno sequences.

Structural Gene--A DNA sequence which encodes through its template or messenger RNA ("mRNA") a sequence of amino acids characteristic of a specific polypeptide.

Transcription--The process of producing mRNA from a structural gene.

Translation--The process of producing a polypeptide from mRNA.

Expression--The process undergone by a structural gene to produce a polypeptide. It is a combination of transcription and translation.

Plasmid--A nonchromosomal double-stranded DNA sequence comprising an intact "replicon" such that the plasmid is replicated in a host cell. When the plasmid is placed within a unicellular organism, the characteristics of that organism may be changed or transformed as a result of the DNA of the plasmid. For example, a plasmid carrying the gene for tetracycline resistance (Tet^R) transforms a cell previously sensitive to tetracycline into one which is resistant to it. A cell transformed by a plasmid is called a "transformant".

Phage or Bacteriophage--Bacterial virus many of which consist of DNA sequences encapsidated in a protein envelope or coat ("capsid").

Cloning Vehicle--A plasmid, phage DNA or other DNA sequence which is able to replicate in a host cell, characterized by one or a small number of endonuclease recognition sites at which such DNA sequences may be cut in a determinable fashion without attendant loss of an

$0_{1/4}$	1.2	0
	1.2	0
	1.0	0*
	1.2	0**
$0_{1/5}$	0.7	0
	0.7	≤0.2*
	1.0	0**
$0_{1/6}$	0.7	0
	1.0	≤0.2*
	0.5	0**
$0_{1/7}$	0.5	0
	1.2	0*
	<0.2	0.5**
$0_{1/8}$	0	1.7*
	<0.2	1.2*
	0	0.7**
	0	1.0**

---

* DBM cellulose paper method

** Nitrocellulose sheets

Therefore, clone $0_{1/8}$ contains a recombinant DNA molecule capable of hybridizing FMDV mRNA from total RNA containing FMDV mRNA. Non-specific RNA-DNA binding is highly unlikely, because a comparison of Fractions 3A and 4A revealed substantially no non-specific binding of STNV DNA in these same experiments. E.g. as monitored by translation in a rabbit reticulocyte lysate in the presence of ^{35}S-methionine, followed by gel electrophoresis, as described above. Clone $0_{1/8}$ was designated E. coli HB101 (C-pBR322(Pst)/HFIF1 ("C-HB101-pHFIF1"), its recombinant DNA molecule C-pBR322 (Pst)HcFIF1 ("pHFIF1") and its hybrid insert "cHFIF1 fragment". This nomenclature indicates that the clone and recombinant DNA molecule originated in Ghent ("C") and comprises plasmid pBR322 containing, at the PstI site HFIF cDNA ("HFIF"), the particular molecule being the first located.

IDENTIFICATION OF CLONES CONTAINING RECOMBINANT DNA-MOLECULES CROSS-HYBRIDIZING TO pHFIF1

pHFIF1, isolated above, was used to screen the library of clones, prepared previously, for bacterial clones containing recombinant DNA molecules having related hybrid DNA inserts, by colony hybridization (M. Grunstein and D.S. Hogness, "A Method For The Isolation Of Cloned DNA's That Contain A Specific Gene", Proc. Natl. Acad. Sci. USA, 72, pp. 3961-3965 (1975)). This method allows rapid identification of related clones by hybridization of a radioactive probe to the DNA of lysed bacterial colonies fixed in nitrocellulose filters.

The library of clones stored in microtiter plates as described above, was replicated on similar size nitrocellulose sheets (0.45 um pore-diameter, Schleicher and Schuel or Millipore), which had been previously boiled to remove detergent, and the sheets placed on LB-agar plates, containing tetracycline at 10 ug/ml. Bacterial colonies were grown overnight at 37°C. Lysis and fixation of the bacteria on the nitrocellulose sheets took place by washing consecutively in 0.5 N NaOH (twice for 7 min), 1 M Tris-HCl (pH 7.6)(7 min), 0.5 M Tris-HCl (pH 7.5) and 1.5 M NaCl (7 min), 3 x SSC (0.15 M NaCl, 0.015 M sodium citrate (pH 7.2)(for 7 min)). After thorough rinsing with ethanol and air drying, the sheets were baked at 80°C for 2 h in vacuo and stored at room temperature.

A Hinf I restriction fragment specific for the pHFIF1 fragment (inf.a) served as the probe for colony hybridization, described infra. This fragment (~170 base-pairs) was purified by electrophoresis of the Hinf digestion products of pHFIF1 in a 6% polyacrylamide gel. After staining the DNA bands with ethidiumbromide, the specific fragment was eluted, reelectrophoresed and ^{32}P-labelled by "nick translation" (P.W.J. Rigby et al., "Labeling Deoxyribonucleic Acid To High Specific Activity In Vitro By Nick Translation With DNA Polymerase I", J. Mol. Biol.,

237-251 (1977)) by incubation in 50 µl 50 mM [illegible] (pH 7.4), 10 mM $MgCl_2$, 20 mM β-mercaptoethanol, [illegible] 2.5 µl each of dCTP, dTTP and dGTP at 400 µM, [illegible] probes α-^{32}P-ATP (Amersham, 2000 Ci/mmole) and 2.5 [illegible] DNA-polymerase I (Boehringer) at 14°C for 45 min. [illegible] unincorporated deoxynucleoside triphosphates were removed [illegible] by gel filtration over Sephadex G-50 in T.E. buffer. The ^{32}P-labelled DNA was precipitated with 0.1 vol of [illegible] 3 M sodium acetate (pH 5.1) and 2.5 vol of ethanol at [illegible]°C.

Hybridization of the above probe to the filter impregnated DNA was carried out essentially as described by D. Hanahan and M. Meselson (personal communication): The filters, prepared above, were preincubated for 2 h at 68°C in 0.1% Ficoll, 0.1% polyvinylpyrrolidone, 0.1% bovine serum albumin, 0.15 M NaCl, 0.03 M Tris-HCl (pH 8), 1 mM EDTA, and rinsed with 0.02% Ficoll, 0.02% polyvinylpyrrolidone, 0.02% bovine serum albumin, 0.75 M NaCl, 0.15 M Tris-HCl (pH 8), 5 mM EDTA and 0.5% SDS. The hybridization proceeded overnight at 68°C in a solution identical to the rinsing solution above using the ^{32}P-labelled probe which had been denatured at 100°C for 5 min prior to use. The hybridized filters were washed twice with 0.3 M NaCl, 0.06 M Tris-HCl (pH 8), 2 mM EDTA for 2 h at 68°C before air drying and autoradiography.

About 1350 clones, originating from the 600-900 DNA size class, were screened. Thirteen colonies, including pHFIF1, gave a positive result. These clones were designated G-N3)01-pHFIF1 to 13 and their recombinant DNA molecules pHFIF1 to 13. One of the clones, pHFIF2, was hybridized with poly(A) mRNA containing F IF mRNA and assayed using DBM-cellulose paper (supra). Because the total IF-RNA activity was detected in the hybridized fraction and the unhybridized RNA did not contain any detectable activity, it is clear that clones identified by colony hybridization to a part of the pHFIF1 fragment also hybridize to F IF mRNA.

It is, of course, evident that this method of clone screening may be employed equally well on other clones containing DNA sequences arising from recombinant DNA technology, synthesis, natural sources or a combination thereof or clones containing DNA sequences related to any of the above DNA sequences by mutation, including single or multiple, base substitutions, insertions, inversions, or deletions. Therefore, such DNA sequences and their identification also fall within this invention. It is also to be understood that DNA sequences, which are not screened by the above DNA sequences, yet which as a result of their arrangement of nucleotides code for those polypeptides coded for by the above DNA sequences also fall within this invention.

CHARACTERISATION OF THE F IF-RELATED RECOMBINANT PLASMIDS

The thirteen clones which were detected by colony hybridization were further characterized. A physical map of the inserts of these clones was constructed and the orientation of the inserts in the various clones was determined.

The physical maps of the plasmids were constructed by digestion with various restriction enzymes (New England Biolabs) in 10 mM Tris-HCl (pH 7.6), 7 mM $MgCl_2$ and 7 mM β-mercaptoethanol at 37°C by well-known procedures. The products of digestion were electrophoresed in 2.2% agarose or 6% polyacrylamide gels in 40 mM Tris-HOAc (pH 7.8), 20 mM EDTA. They were analyzed after visualization by staining with ethidiumbromide and compared with the detailed physical map of pBR322 (J.G. Sutcliffe, supra). Restriction maps of the different plasmids were constructed on the basis of these digestion patterns. These were refined by sequencing the DNA inserts in various of the plasmids, substantially by the procedure of A.M. Maxam and W. Gilbert "A New Method For Sequencing DNA", Proc. Natl. Acad. Sci. USA, 74, pp. 560-564 (1977).

Plasmid DNA was prepared from various of the pHIF1-13 in accordance with this invention by the method

The comparison of the first 13 amino acids of authentic fibroblast interferon (Knight et al., supra) and the sequence deduced from the composite gene of Fig. 4 show no differences. The amino acid compositions determined directly for authentic fibroblast interferon on the one hand and that deduced from the sequence of the composite gene of this invention on the other also show substantial similarities. Fig. 6 displays a comparison of these compositions.

Although none of the recombinant DNA molecules prepared in accordance with this invention contain the complete DNA sequence for fibroblast interferon, a combination of portions of the inserts of these recombinant DNA molecules to afford the complete F IF DNA gene sequence is within the skill of the art. For example, by reference to Fig. 5, it can readily be seen that the PstI-BglII fragment of pHFIF6 could be joined with the PstI-HaeII fragment of pHFIF7 or the EcoRI-PstI fragment of pHFIF6 could be joined with the PstI-HaeII fragment of pHFIF7 or the BglII-PstI fragment of pHFIF6 could be joined with the PstI-BglII fragment of clone 7 to form the composite F IF gene. The joining of these fragments could be done before or after insertion into a desired plasmid.

Micro-organisms and recombinant DNA molecules prepared by the processes described herein are exemplified by cultures deposited in the culture collection Deutsche Sammlung von Mikroorganism in Gottingen, West Germany on April 2, 1980, and identified as HFIF-A to C:

A: E. coli HB101 (G-pBR322(Pst)/HFIF3)

B: E. coli HB101 (G-pBR322(Pst)/HFIF6)

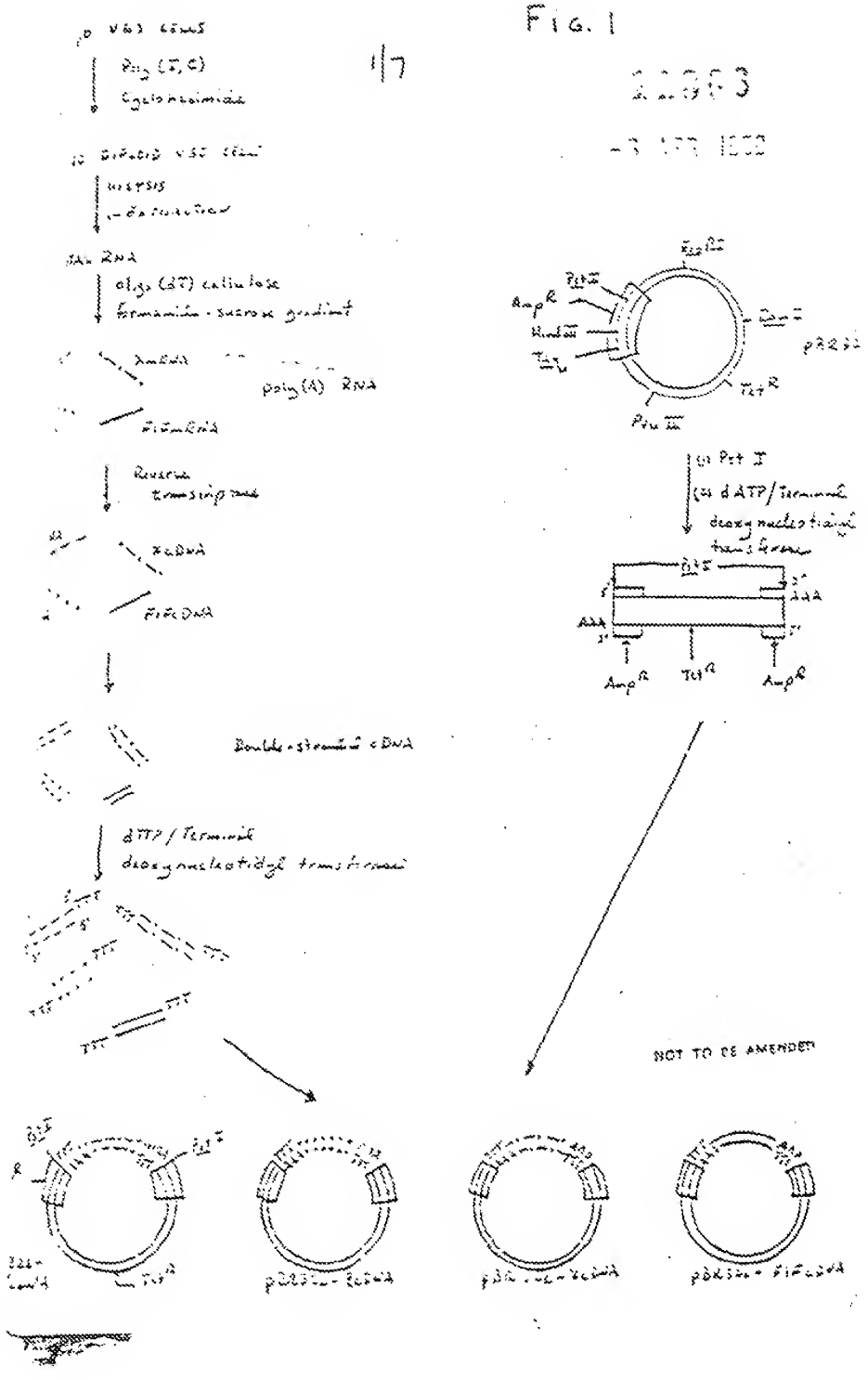
C: E. coli HB101 (G-pBR322(Psr)/HFIF7)

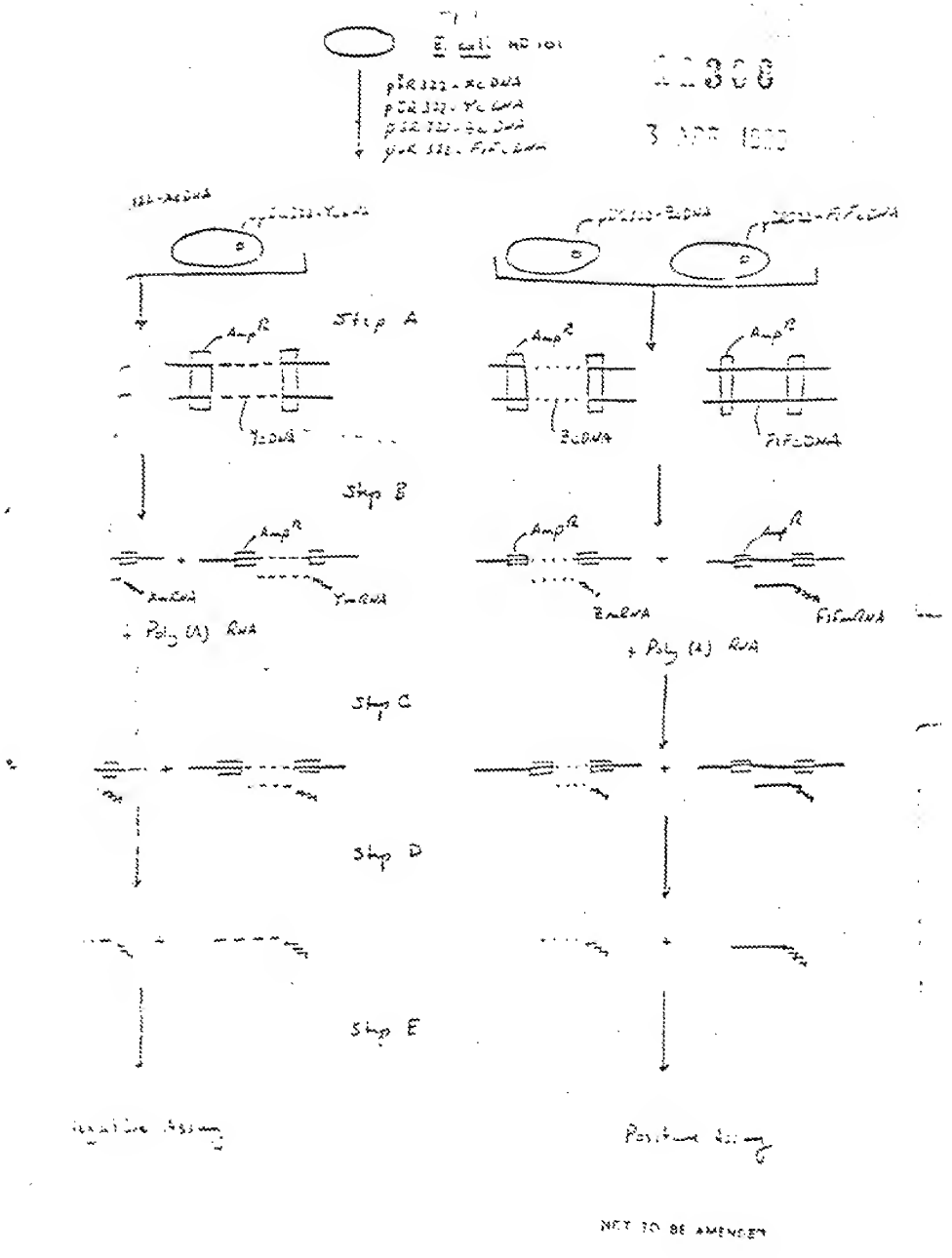
These cultures were assigned accession numbers DSM 1791-1793, respectively.

While we have herein before presented a number of embodiments of this invention, it is apparent that our basic construction can be altered to provide other embodiments which utilize the processes and compositions of this invention. Therefore, it will be appreciated that the

FIG. 1

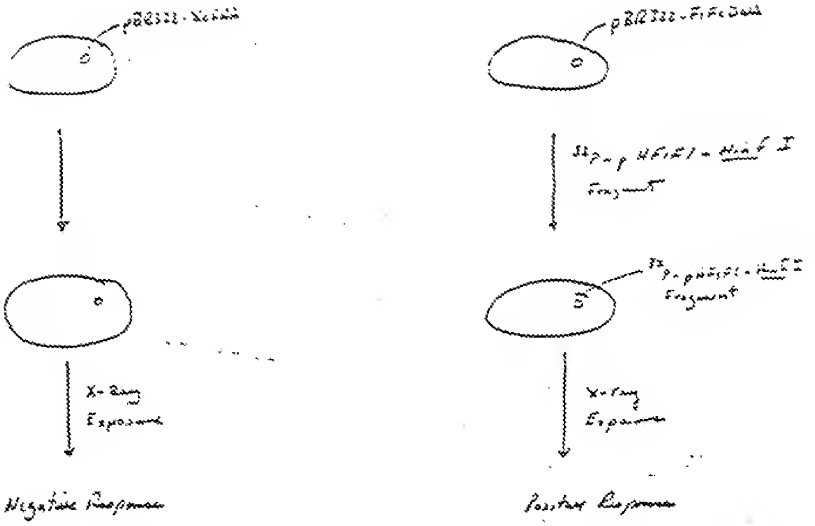
1/7





NOT TO BE AMENDED

Fig. 3 3/7



6/7

Amino Acid Composition of Human Fibrinogen Immunogen

Amino Acid	Composition from direct analysis by Tan et al.	from direct analysis by Knight et al.	deduced from nucleotide sequence	
Asp	20.6	18.9	5	17
Asn			12	
Thr	8.0	6.8	7	
Ser	11.7	10.5	9	
Glu	27.6	27.0	13	24
Gln			11	
Pro	4.4	2.7	1	
Gly	5.4	7.8	6	
Ala	9.3	10.0	6	
Cys	N.D.	1.7	3	
Val	7.9	6.0	5	
Met	trace	2.9	4	
Ile	10.0	9.0	11	
Leu	26.9	29.4	24	
Tyr	3.2	7.5	10	
Phe	7.7	9.4	9	
His	4.6	4.9	5	
Lys	12.3	11.6	11	
Arg	8.6	10.9	11	
Trp	0.0	1.0	3	
TOTAL	168	169	166	

Fig. 6

NOT TO BE AMENDED

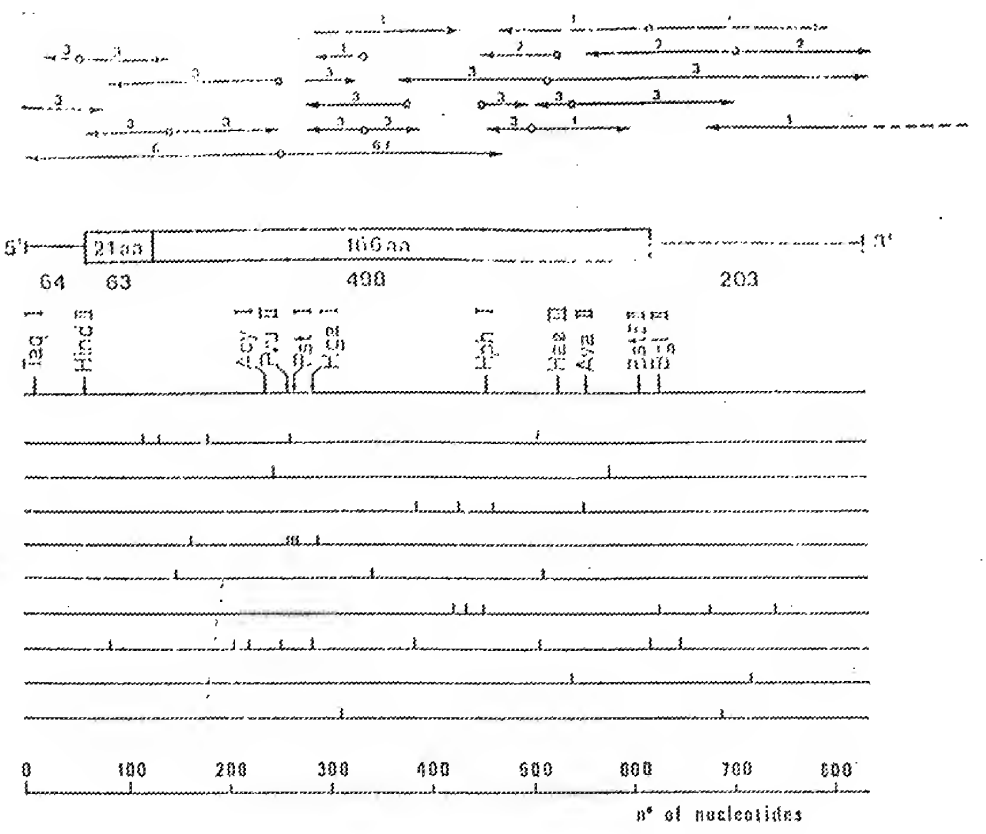


FIG. 7

1

Interferon clones

Uitgaande van Clone 3

informatie uitgeknipt met Taq en het fragment
pBR322 gedeeltelijk
GC getailed in de Pst site van SRK2311

Clones SRK I3-4 : Sense orientatie
SRK I3-3 : Nonsense orientatie

Testen of aktiviteit : negatief.

Uitgaande van kombinatie van Clone 6 met Clone 7

A) EcoR1 – Pst uit clone 6 } van SRK2311
Pst – HaeII uit clone 7 }

1 Clone: pPLa HFIF 67-1
(gemaakt in K12 [illegible] en overgebracht
naar W[illegible] (+ dan later analyse bevestigd))
ook naar M5219

B) BglII – Pst uit clone 6 } van Bam site
Pst – BglII uit clone 7 } van [illegible] of ST28 / STL24

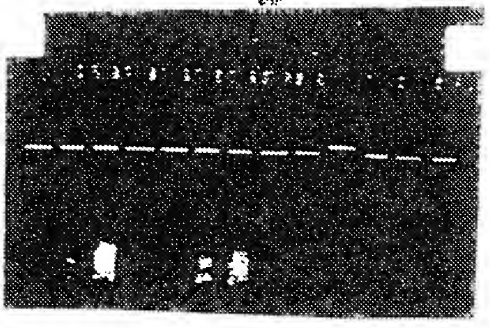
Dec 30 02

[redacted]
In presence of examiner
en présence de l'examinateur

This is EXHIBIT FIERS-20
to
the Affidavit of Walter C. Fiers
sworn before me
this 19 th day of November, 2001

Commissioner for Oath or Notary Public

12 clone wt mfg pSTL24 Eglin 1-7



P3C

Hin II

Interferon clones

Uitgaande van Clone 3

informatie uitgebracht met Tac in het plasmide pBR322 gedeelte
EC getailed in de Pst site van SRK1311

clones SRK I3-4 : Sense orientatie
SRK I3-3 : Nonsense orientatie

Testen op activiteit : negatief.

Uitgaande van kombinatie van Clone 6 met Clone 7

A) EcoR1 - Pst uit clone 6 } naar SRK2311
Pst - Hae II uit clone 7 }

1 Clone: pPla HFIF (7-1
(gemaakt in K12 ΔH1 (λ) en overgebracht naar M F1 (→ dan door analyse bevestigd))
ook naar M 5219

B) BglII - Pst uit clone 6 } van Bam site
Pst - BglII uit clone 7 } van of ST98 / STL24

Dec 30 .02

In presence of examiner
en présence de l'examinateur

This is EXHIBIT FIERS-20
to
the Affidavit of Walter C. Fiers
sworn before me
this 19 th day of November, 2001

Commissioner for Oath or Notary Public

Detail van konstruktie B.

2 µg Cl 6
2 µg Cl 7 / geknipt met PstI x BglII

geligeerd 4 uur 25°C in aanwezigheid van BglII

helft naar 1 µg Bam geknipt ~~...~~ pST28
helft naar 5 µl Cl. bm. Bam geknipt pSTL24

geligeerd in aanwezigheid van Bam x BglII
16 uur 25°C
nog nageknipt met Bam

telkens helft getransformeerd naar { K12 $r^- m^-$ (λ) / NF_1

Transformanten		# totaal
NF_1 met ST28 BglII 1-7	:	300
NF_1 met STL24 BglII 1-7	:	50
K12 met ST28 BglII 1-7		1000
K12 met STL24 BglII 1-7		500

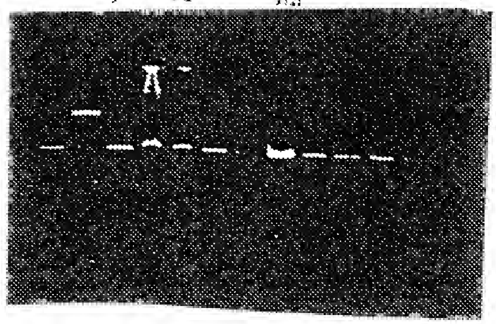
[redacted]

Clone Analyse van transformanten

Mini SDS methode

12 clones uit RF_1 met ST28 BglII 6-7

1 2 3 4 5 6 7 ST22 8 9 10 11 12

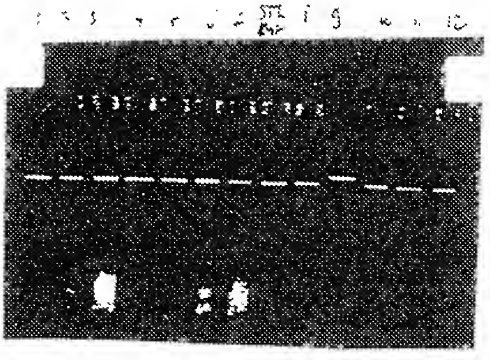


Pst digest

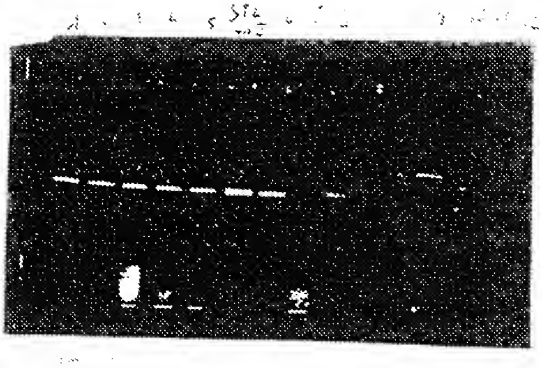
1 2 3 4 5 6 7 ST22 8 9 10 11 12

Hin 2 digest

12 clones with neg pSTL26 Sph I-7



Pst



Hin III

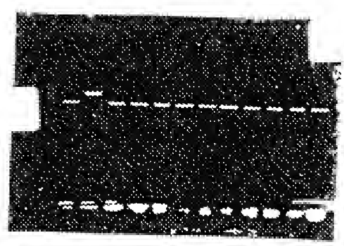
ST mutanten: 11 originele ST22 clones
1 heeft 2 PstI sites maar is wel te groot
heeft 1 Hind III fragmenten

ST2 deleties: 11 originele clones
1 is iets groter maar heeft slechts 1 PstI site

<u>Verder werk met ST2 deleties</u>

20 kolonies werden gekweekt in 1 ml 10 mM Tris
12 opstellingen
daaruit werd DNA geisoleerd en gedigereerd met
Bam → Transformatie van ± 10⁵ ... (λ)
Per opstelling werd 1 transformant geanalyseerd
door ... methode

Restricties met PstI



→ geen enkele clone heeft 2 PstI sites
→ op 1 na zijn ze alle originele ST29